INDUSTRIAL CHEMISTRY OF

Colloidal & Amorphous

MATERIALS

INDUSTRIAL

CHEMISTRY OF

Colloidal and

Amorphous

MATERIALS

B Y

WARREN K. LEWIS, Ph.D., *Professor in the Department of Chemical Engineering, Massachusetts Institute of Technology*

LOMBARD SQUIRES, S M., *Chemical Engineer with E. I. du Pont de Nemours Company, Wilmington, Delaware*

GEOFFREY BROUGHTON, Sc.D., A.I.C., *Chemical Engineer with the Eastman Kodak Company, Rochester, N. Y.*

NEW YORK · THE MACMILLAN COMPANY

To
WILLIAM H. WALKER
Pioneer in Chemical Engineering Whose
Teaching Furnished the Inspiration for This Book

Preface

This book is a text on the industrial chemistry of colloidal and amorphous materials. Its purpose is to give readers starting study of the subject with a general background of chemistry a satisfactory introduction to the field. It is not intended to be merely descriptive of the methods of the industries but to give the reader such an insight into the underlying phenomena that he can appreciate the potentialities and limitations of processes and materials and develop his capacity for controlling and using them. This presupposes a mastery of those phases of science underlying the industry, a fact which involves two serious difficulties.

In the first place, whereas the chemist starting the study of other fields of applied chemistry, such as so-called heavy chemical industry, brings an adequate background of training in inorganic, organic, analytical, and physical chemistry and the allied sciences, the experience of many years, including work with graduate students from a wide variety of institutions, demonstrates that this is not true in the colloidal and amorphous fields. The deficiency is mainly in physical chemistry, for the reason that its field is so broad that, to be thorough, elementary instruction must be restricted in scope. In choosing the subject matter to be included, instructors in physical chemistry have, in general, omitted or at least not emphasized much of utmost importance to an adequate understanding of the type of problems with which this text is dealing. Because of the conviction that an adequate scientific background is the only proper foundation for the study of industrial technology, it has seemed wise to devote the first part of this book to the presentation of the basic material lacked by the average reader.

The subject matter thus included is not new. There are many excellent texts on colloidal chemistry and on plastic materials. The best of these, however, fail to meet the needs of the situation in two respects: they underemphasize or omit parts of the subject which, while elementary, are essential to an understanding of many phenomena of outstanding industrial importance and they fail to make clear the inter-relationships of the colloidal and amorphous states. These introductory chapters are intended not so much to cover the whole field of the physical chemistry of colloidal and amorphous materials as to emphasize those phases of major importance in industry or which seem likely to become important in the not too distant future. Attention is focused on the nature and effects of inter-molecular forces and on the structure of surfaces of discontinuity. It is then attempted to use these concepts constructively in the interpretation of the complicated phenomena encountered in colloidal and amorphous systems.

The second difficulty complicating presentation of the subject is caused by its complexity. In a specific problem knowledge of the facts is often inadequate and occasionally in dispute, while understanding of underlying causes is limited and uncertain. Experts disagree on the explanations of phenomena of fundamental importance. In such a situation the scientist tends to withhold judgment but the technologist is faced with the necessity of making decisions and taking action. Many individuals prefer a purely empirical approach, but it is the conviction of the writers that it is best to submit each problem to a careful, theoretical analysis, even though in the nature of the specific case this may be known to be inadequate. A major purpose, therefore, is the introduction of the student to such a method of attack.

Progress in this field has been achieved not so much by the intensive study of a limited number of facts as by the effort to interpret each fact in the light of all others related to it. To do this effectively the first essential step is to differentiate clearly between the facts and phenomena on the one hand and the opinions, interpretations and conclusions based upon them. In the literature these are not infrequently confused. To aid the

student it has seemed essential to present the facts of the given situation ahead of and segregated from the discussion of it. This results in clumsiness of presentation and a certain amount of repetition, but the price is necessary to secure clarity of analysis. It is more important for the student to have opinions of his own and to know the facts on which they are based than to accept the opinions of the authors. Only thus will he develop capacity to revise and readjust his opinions as new data throw new light on the problems.

There are few fields of the physical sciences in which it is more necessary to maintain breadth of view. For example, concepts of the structure of rubber have developed rapidly and secured wide acceptance during the last few years. However, this has been due not so much to new data on rubber itself, much of the information concerning which was available a quarter of a century ago, as to new facts on other relatively unrelated materials, such as the synthetic polymers. Again, light is thrown on one industry by data on others quite divergent. To help the reader appreciate these complex interrelationships, cross-references are used extensively throughout the text. Furthermore, the brief bibliographies following the chapters are intended to give an introduction to the literature that will emphasize the breadth of the various fields and the differences in approach.

The group of industries here covered are in a stage of rapid development. This tends to diversify the practices of different plants far more than would be characteristic of a stabilized field. Detailed descriptions of parallel processes in different plants are undesirable and generalizations as to manufacturing methods are unsafe and may be misleading. The reader must keep in mind that the descriptions given of processes and products are intended to be typical rather than specific.

This book is the outgrowth of personal, practical experience which goes back over thirty years. It embodies the results of twenty-five years teaching the subject at the Massachusetts Institute of Technology. The text has been prepared entirely by members of its staff, all but one of whom have since gone into industry. Over this long period many professional associates

have through suggestions, discussion and information made contributions of major value which are deeply appreciated. The number of these individuals is so large that acknowledgment of our obligations must be made collectively. Over a dozen years ago, Dr. P. K. Frolich started work on the preparation of the manuscript. His responsibilities in industry made it necessary for him to discontinue, but, despite the changes necessitated by the developments of the intervening years, evidences of his contributions can be found in various parts of the text.

<div align="right">

WARREN K. LEWIS
LOMBARD SQUIRES
GEOFFREY BROUGHTON

</div>

December, 1941

Table of Contents

INDUSTRIAL CHEMISTRY OF
Colloidal & Amorphous
MATERIALS

Chapter I

Structure of Liquids

For the sake of convenience it is usual to classify matter into three states—gaseous, liquid, and solid. A convincing theoretical picture of the first of these was given in the middle of the last century by the kinetic theory of gases; of late years, X-ray investigations have led to a satisfactory concept of the structure of solids. On the other hand, the internal structure of liquids still remains far from clear; hence, it is the purpose of this chapter to analyze their behavior, because of the light it throws on the properties of matter in two of its most important forms, the colloidal and the amorphous states.

The Kinetic Theory and the Structure of Liquids.

The kinetic theory pictures a gas as a group of individual molecules, simultaneously engaged in rapid and chaotic motion. In a perfect gas each molecule is assumed to possess on the average a certain kinetic energy of translation, the total of which remains constant after collision; *i.e.*, to act on collision as though perfectly elastic.* Naturally, at any one instant the molecules will be moving at various speeds, the velocity distribution being of the type shown graphically in Fig. 1. The average kinetic energy of translation of the molecules is dependent upon the temperature, and independent of molecular weight and other
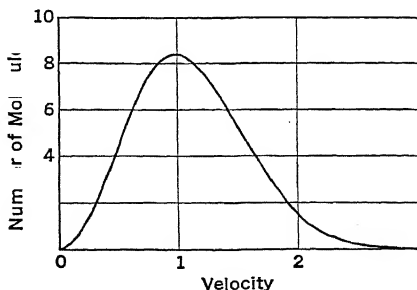

FIG. 1. Molecular Velocity Distribution of Gases.

* The kinetic energy may be of three major types, translational, rotational, or oscillatory. Momentarily, by the collision, some of the translational energy may be converted into other forms, but this is later reconverted to translational energy.

1

factors; it is equal to $3RT/2$, R being the gas constant and T the absolute temperature.* As a result of the rebounds on collision, the molecules tend to separate and occupy a greater volume; this tendency in turn is governed by the average translational energy of the molecules. Opposing this attempt to expand is the external pressure on the containing vessel. There is, therefore, a balance between a disgregating force, resulting from the thermal agitation of the molecules, and an aggregating force, the external pressure. For given conditions the gas comes to an equilibrium corresponding to this balance between the opposing forces. When these concepts are thrown into quantitative form, they readily serve to derive the gas law equation, $PV = nRT$. For a given temperature, the molal concentration of a gas varies directly with the pressure, because the disgregating force in the gas, due to the collisions of the molecules brought about by their thermal agitation, is proportional to the number of these collisions, and hence to the concentration of the gas, while the counterbalancing aggregating force is the external pressure.

A liquid behaves far differently, molal concentration or density remaining relatively constant over a wide range of pressure, so that the application of the kinetic theory to a pure liquid seems at first to lead to insuperable difficulties. Thus, liquid water at 0° C., existing under its own vapor pressure of 4.6 mm., has a density of 0.9999 gms. per c.cm., over 200,000-fold that of the saturated vapor in equilibrium with it. If the pressure be increased isothermally to 100 atmospheres, the density of the water increases to 1.005, *i.e.*, by 0.5%. The water vapor in equilibrium with the liquid at this temperature is, of course, physically incapable of undergoing a large increase in pressure without condensation, but, could it likewise be compressed to 100 atmospheres as a vapor, the gas laws would indicate an increase in density of 16,500-fold, in contrast to the inconsequential change in density of the liquid. There must be some significant factor, relatively unimportant in the vapor state but dominant in the liquid.

---

* For the monatomic gases, *i.e.*, certain metallic vapors, the noble gases, helium, argon, etc., this will also be their total kinetic energy, since their only possible form of kinetic energy is translational.

Were the slowing down of the molecules in the liquid state a major factor, it should account for the energy changes involved in condensation. It is easily shown that the total energy of translation of the vapor molecules is insufficient to account for more than a small fraction of the energy released on condensation to the liquid state.* Moreover, the kinetic theory leads to the conclusion that the molecule in the liquid must have an effective kinetic energy of translation identical with that of the vapor molecule at the same temperature, an energy which therefore is theoretically independent of molecular weight and structure, and a function of the temperature only.† Another possible explanation is the compressibility of the molecule itself, but, as explained elsewhere, the evidence indicates that this factor, important at high pressures, is insignificant in the phenomenon of liquefaction at ordinary pressures. It seems beyond doubt that the explanation of the underlying mechanism of liquefaction can be found only by assuming the existence of large forces of intermolecular attraction, first postulated by Young and Laplace in the early years of the nineteenth century.

Internal Pressure.

In Fig. 2, let CD represent the interfacial surface between a pure liquid A, and its vapor B, while E is a single molecule in the surface. On the liquid side of E, because of the high density of the liquid, there is a large number of molecules; on the vapor side, how-

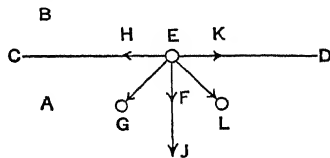FIG. 2.

* Thus, in the following table the internal energies of vaporization and kinetic energies of translation (3RT/2) of a number of substances at their boiling points are given:

Substance	Boiling Point, ° K	Internal Latent Heat, cals./gm. mol	Total K.E., cals./gm. mol	Ratio
Argon	87	1286	259	4.8
Water	373	8960	1111	8.1
Hydrogen chloride	188	3477	560	6.2
Benzene	353	6650	1052	6.3

It is seen that in every case, even for the monatomic gas argon where the translational kinetic energy is also the total kinetic energy, the internal latent heat of vaporization is many times the total translational kinetic energy of the molecules.

† Evidence substantiating this conclusion will be referred to on p. 13.

ever, the number of molecules is small and may be negligible. A molecule in the liquid, such as G, will attract E with a force, represented in magnitude and direction by EG, which may be resolved into a horizontal component EH and a vertical one EF. Another molecule in the liquid, L, symmetrically placed with reference to G on the opposite side of EJ, will similarly attract E with a force represented by EL, resolvable into components EK and EF. The horizontal components, EK and EH, acting in opposite directions, neutralize each other, but the vertical components are additive. Thus, the molecules on the liquid side of E attract it with forces, the components of which parallel to the surface neutralize each other, whereas the inwardly directed components perpendicular to the surface are additive. In consequence, E is attracted toward the interior of the liquid by a total unbalanced force EJ. Since this is true of every molecule in the surface, there is a force acting inwardly over the whole surface. This result is equivalent to subjecting the liquid to a pressure, caused, not by an imposed force, but by the mutual attractions of the molecules themselves. This pressure is the so-called internal pressure of the liquid. It may be visualized as that pressure which would have to be imposed upon the bulk of the liquid, in order to reduce its volume to the actual value against the disgregating forces, *e.g.*, of thermal agitation, tending to keep the molecules apart, were there no attractive forces at all operating between the molecules. In the light of the gas laws and the high densities of liquids, it is clear that the internal pressures of liquids must be large in magnitude.

Each molecule in the interface CD, Fig. 2, is moving in disordered motion, the statistical distribution of velocities being that shown in Fig. 1. Assume a molecule fortuitously to acquire a velocity with a component outwardly directed from the liquid. As it starts away from the interface it is attracted backward by the intermolecular forces, and, in general, it will be slowed up, its direction reversed, and it will return finally into the surface. If, however, its outwardly directed component of velocity happens to be sufficiently large, the molecule will be able to continue its motion into the vapor despite the retarding effect of the

molecular attraction. Since the distribution of relative veloci-
ties (Fig. 1) is independent of the type of molecule involved,
comparing any two molecular species at the same temperature,
the relative chance of escape is determined primarily by the
respective molecular attractions. At equilibrium, rate of escape
equals that of return. At low pressures the latter is clearly
proportional to the vapor pressure. Hence, the equilibrium
vapor pressure of the liquid is an inverse measure of the attrac-
tive forces.

Keeping this analysis in mind, the experimentally determined
characteristics of vapor pressure throw light on the nature of
the intermolecular forces. It is well recognized that the vapor
pressure is independent of the mass of liquid under considera-
tion.* Thus, the vapor pressure of a liquid is independent of
its depth. This can be true only in case molecules in the liquid
at a considerable distance from the vapor-liquid interface exert
no appreciable attraction upon one in the surface. In other
words, it appears that only molecules in the immediate neigh-
borhood of the one under consideration can act upon it to a
significant degree. This is equivalent to assuming a limit of
distance beyond which intermolecular attractions become negli-
gible. Only molecules within a sphere whose radius is this
distance, the so-called sphere of action of the intermolecular
forces, can be effective. In other words, these forces are not
gravitational in type.

Visualize a molecule C in the interface
AB between a liquid and its vapor (Fig. 3).
Of the liquid CD, lying between two cones
apexed at C, of half-angle θ and $\theta + d\theta$
respectively, $d\theta$ being small, consider the
differential amount of liquid in this zone
between the concentric spherical surfaces
centered at C with radii of r and $r + dr$,
respectively. The area of the differential
spherical surface is obviously proportional
to r^2, and the differential volume of liquid

Fig. 3.

under consideration, to $r^2\, dr$. If one can assume the attraction of this
liquid on the molecule C proportional to its mass and to some inverse power
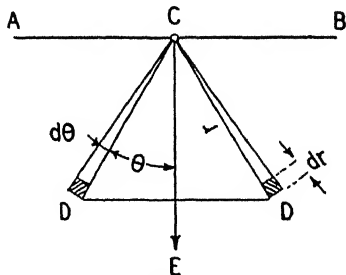of the distance $1/r^n$, the component df, directed inwardly normal to the

* See, however, pp. 87 and 110.

liquid surface, of the attractive force in the direction CD must be

$$df = \alpha r^2 \cos\theta dr \frac{d\theta}{r^n} = \alpha \cos\theta d\theta r^{2-n} dr,$$

where α is a proportionality constant, the value of which depends on the case in hand. Integrating, the total, internally directed force on C is

$$f = \frac{}{3 - n} \quad \left. \vphantom{\frac{}{}} \right]_{r=\infty}^{r=a}.$$

The limits are obviously the average effective distance a, from C to the nearest molecule in the direction CD, and infinity. The experimental evidence indicates, however, that this attractive force is uninfluenced by the thickness of the liquid layer (*i.e.*, by r) beyond quite small values. Thus the vapor pressure of a liquid, which must certainly be modified by any change in the surface attractions, is unaffected by liquid depth, provided the surface is flat, except in those cases in which the layer is very thin. This is possible only if all effect on the integral disappears at large values of r; this requires that n be greater than 3. Hence, the intermolecular forces are not gravitational, as this would require that n be 2.

The argument just outlined is based on certain indefensible assumptions. Probably the most serious is the implication that liquid density remains constant into the interface itself. By analogous reasoning, which, however, because it endeavors to allow for these complicating factors, becomes more involved, Edser proved, still arbitrarily assuming the validity of an inverse power law of attraction, that the value of n must exceed 4, and, on the basis of the properties of liquids, concluded that n is of the order of magnitude of 7 to 9.

It is seen therefore that, since the internal pressures of liquids are large and the spheres of action of the molecules extremely small, the attractive forces between molecules must be very large, although operating only over small distances. However, at high pressures the compressibility of liquids is small, far smaller than at ordinary pressures.* This shows that at high

* This is indicated by data of Bridgman [*Proc. Am. Acad. Sci.* 49, No. 1, 4 (1913)] on the compressibilities of various liquids, of which the following is an example.

COMPRESSIBILITY OF CARBON DISULFIDE AT 20° C.

Pressure in kg./cm.²	Volume	Decrease in Volume
1	1.0235	
2,000	0.9173	0.1062
4,000	0.8647	0.0526
6,000	0.8295	0.0352
8,000	0.8022	0.0273
10,000	0.7805	0.0217
12,000	0.7638	0.0167

pressures, when the molecules are very close together, there must be a considerable repulsive force acting between them.*

The Estimation of Internal Pressure.

So far no technique has been developed for the direct measurement of internal pressure, or for its calculation by methods not open to serious criticism.

Two methods of estimating internal pressure may be mentioned. Referring again to Fig. 2, consider two bodies of liquid at the same temperature, in which the density in the first case is double that in the second. Assuming completely random distribution of the molecules in both cases, in the second a given molecule E in the interface will be attracted inwardly by only half as many molecules as in the first. Since these attracting molecules are distributed throughout half the sphere of action, the attractive force in any given direction is halved in the second case, so that the total inward attraction EF is also halved. However, halving the density will presumably cause the total number of molecules in the interfacial surface to be halved, so that the total inward force per unit area, $i.e.$, the internal pressure, is reduced to one-fourth its original value. This is equivalent to writing

$$\pi = a\rho^2 \tag{1}$$

where π is the internal pressure, a a constant, specific for a given liquid, and ρ the density of the liquid. At constant temperature, in order to separate the molecules by expanding the fluid, energy must be supplied to overcome the net internal attractions of the fluid molecules, $i.e.$, the internal pressure. This "disgregation work" must, therefore, be

$$dW = \pi dV = a\frac{dV}{V^2} \tag{2}$$

or integrating,

$$W = a\,(\rho_1 - \rho_2). \tag{3}$$

* This is understandable, since according to modern theories of atomic structure the atom is pictured as an assemblage of electrons grouped around a positive nucleus. When two atoms come sufficiently close together, their outer shells of electrons are deformed, and as a consequence powerful repulsive forces are set up. These repulsive forces must also be responsible for the rebound and apparently elastic behavior of molecules upon collision (see p. 27).

Assuming that all the energy of vaporization, other than that expended in doing external work, is consumed in overcoming the intermolecular attractions, this expression allows a to be evaluated by equating W to the internal energy of vaporization. The internal pressure, if the assumption of substantially uniform distribution of the molecules is correct, is then given by (1).

The method of calculating internal pressure, the objections to which are perhaps least serious, is that based on the thermodynamic equation of state of the liquid. The relations between the external pressure P, the volume V, the absolute temperature T, and the total energy E, represented by the expression,

$$P + \left(\frac{\partial E}{\partial V}\right)_T = T\left(\frac{\partial P}{\partial T}\right)_V, \tag{4}$$

are a necessary consequence of the second law of thermodynamics. The term $(\partial E/\partial V)_T$ represents the energy absorbed in expanding the liquid at constant temperature, and, since the effective kinetic energy of the molecules should not change if the temperature is kept constant, this term includes the energy used in overcoming the forces of molecular attraction. The energy necessary to increase the volume at constant temperature is $(\partial E/\partial V)_T dV$. Granting that this is also equal to πdV, $(\partial E/\partial V)_T dV = \pi dV$, or $\pi = (\partial E/\partial V)_T$. Thus the term $(\partial E/\partial V)_T$ is, in fact, equal to the internal pressure, provided no other form of energy is consumed on isothermal expansion. Since this pressure is frequently of order of magnitude greater than the external pressure P, it is substantially equal to the right hand member of the equation, which thus also becomes a measure of internal pressure. Its use is generally more convenient, since the change in internal energy of a liquid with volume is difficult to measure. $T(\partial P/\partial T)_V$ can be calculated directly from two easily measurable physical properties, the coefficient of thermal expansion, $\alpha = 1/V\,(\partial V/\partial T)_P$, and the coefficient of compressibility of the liquid, $\beta = -\,1/V\,(\partial V/\partial P)_T$.

Since $(\partial P/\partial T)_V = -\,(\partial V/\partial T)_P/(\partial V/\partial P)_T$, substituting the result in the right hand member of (4), gives

$$\pi = T\left(\frac{\partial P}{\partial T}\right)_V = -\,T\frac{\alpha}{\beta} \tag{5}$$

Thus, the internal pressure π is equal to the ratio of the coefficient of thermal expansion to the coefficient of compressibility, times the absolute temperature. Since β varies with the pressure, the value to be substituted in this equation should be determined at the same pressure as the expansion measurements.

Although the agreement between values calculated by different methods is poor, there is no doubt that for liquids internal pressure is very high (Table I), probably of the order of magnitude of a few thousand atmospheres for organic liquids and probably well over 10,000 atmospheres for such liquids as water and mercury.

TABLE I. INTERNAL PRESSURE π, IN ATMOSPHERES
AT 20° C.

Substance	From Latent Heat	From α and β
Pentane	1,700	2,800
Carbon tetrachloride	2,600	3,700
Chloroform	3,570	3,900
Benzene	3,300	3,900
Bromine	5,600	5,100
Mercury	36,000	14,000
Water	23,000	1,200 [a]

[a] Had this value for water been computed at its point of maximum density, 4° C., it is obvious that the result would be zero. The formulas for internal pressure break down completely in the face of abnormalities in liquid behavior, such as those in water, a fact that keeps one skeptical as to their validity even in other cases.

The Nature of the Attractive Forces.

The hypothesis of the attractive forces within a liquid as the aggregating power responsible for liquefaction, first suggested by Young and Laplace and developed principally by van der Waals and his school on physical grounds, long met with little recognition from chemists. However, the growing knowledge of atomic and molecular structure and the development of wave mechanics have led to an increasingly clear understanding of the nature of intermolecular forces and the mechanism of their action. They are undoubtedly due to stray fields of force around the surface of the molecule, which are again due to the electronic structures of the atoms comprising the molecule. While the normal molecule is always completely neutral electrically, in the sense that the positive and negative

charges within it are exactly equal in magnitude, these charges
are not identical in location. The negative electrons are con-
ceived as moving around the exterior of the atom in a series of
concentric shells, while the positive charge is concentrated in
the nucleus. This structure cannot fail to result in stray fields
of force extending outside the molecule itself. In other words,
a surface localization of charges may occur, although the mole-
cule as a whole is electrically neutral. A simple analogy may
serve to make this clear. Two magnets of equal pole strength,
arranged as shown in Fig. 4, so that the over-
all magnetic neutralization for the pair is com-
plete, will yet show a residual, stray field of
magnetic force, easily demonstrated by the
familiar iron filings experiment, in which the
filings arrange themselves along the magnetic
lines. The strength of such a field decreases very rapidly
as distance from the magnetic pair increases. Similarly in
the molecule, although it has no net negative or positive
electrical charge and its charge arrangement is very different,
there are stray fields of electric force. This explains the pos-
sibility of the existence of enormous attractive forces in the
immediate neighborhood of the molecule, which must, be-
cause of their very character, fade away to negligible values as
the distance is relatively slightly increased. Changes in the
chemical constitution of the molecule cannot fail to cause varia-
tions in the location and strength of the surface fields, thus
accounting for the marked influence of molecular structure on
physical properties. The assumption of these fields gives a clear
qualitative explanation of the characteristics of the intermolecu-
lar forces, originally deduced on purely physical grounds decades
before present theories of atomic and molecular structure were
developed.

The liquid state is, therefore, entirely analogous to the gase-
ous state, except that, owing to the higher density, the molecules
are never outside their mutual spheres of attraction. For any
given molecule in a liquid there will exist a neutral zone, in
which the attractions of its neighbors on it will exactly counter-
balance each other. As a molecule passes out of this zone it

FIG. 4.

will speed up, since the attractive forces upon it will become greater. Hence the true, average velocity of the molecule will be greater than its velocity as it passes through the neutral zone. Nevertheless, the translational velocity effective as a disgregating force must be the average velocity in the neutral zone. The difference between this effective velocity and the true average velocity is far greater in the liquid than in the gas, because in the former the molecule is in the neutral zone an inconsequential percentage of the time, in contradistinction to the large percentage of neutral travel in the gas or vapor state.

Exactly the same forces, therefore, operate in a gas as in a liquid, with the intermolecular attractions small in the former but predominant in the latter. The molecules move in the same manner, but with a far smaller free path in the liquid. Moreover, it is easy to show from experimental data that molecular attractions are not absent in the gaseous state; indeed, they are a major cause of the deviations from the gas laws.

Intermolecular Attractions and the Gaseous State.

It is well known that even at moderate pressures real gases do not obey the simple gas laws. In Fig. 5 are given isothermal (0° C.) values of PV/RT for methane at a number of different pressures. Any correction for the volume occupied by the gas molecules should make the term PV/RT greater than unity. Actually it is less, except at very high pressures. The deviation is readily explained by assuming an aggregating force in addition to the external pressure, caused by the

FIG. 5. Compressibility Factors for Methane at 0° C.*

attractive forces between the molecules themselves. At high pressures there will be an opposing effect due to the volumes of the molecules, but at intermediate pressures the attractive forces are more important. These two factors, by which deviations from the gas law can be explained, were allowed
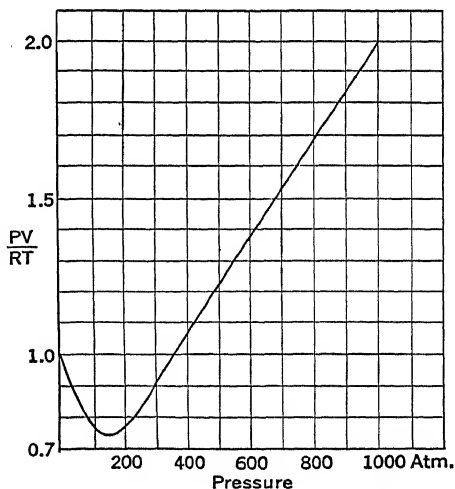
* Data of KVALNES, H. M., and GADDY, V. L., *J. Am. Chem. Soc.* **53**, 394 (1931).

for by van der Waals in his familiar equation:

$$P + \frac{a}{V^2} \ (V - b) = RT.$$

In this, the term a/V^2 is a correction factor for the aggregating force due to molecular attractions (see pp. 4–5) and the term b corrects for the volume of the molecules themselves. Quantitatively the equation is only an approximation, but, used qualitatively, it is often helpful.

Additional evidence for molecular attraction in the gaseous state is provided by the Joule-Thomson effect, the cooling which occurs when gases are allowed to expand through a porous plug so that there is negligible velocity change. The absorption of heat indicates that work is done, *i.e.*, the energy required to overcome the attraction of the molecules for each other.

In both gases and liquids the external pressure and intermolecular attractions act as aggregating forces, the disgregating forces being the kinetic motion of thermal agitation and the relative incompressibility of the molecules themselves. In gases the external pressure is the more important aggregating force, and in liquids, the intermolecular attractions.

Osmotic Pressure.

The assumption of large intermolecular forces of attraction within liquids gives a clearer understanding of the nature of osmotic pressure. This is sometimes visualized, not as a true pressure, but rather as a potential pressure-producing capacity. This is due to failure to visualize the nature of the forces which oppose osmotic pressure and restrain the expansive tendency of the solute. The osmotic pressure operates outwardly as an expansive or disgregating force, and is counterbalanced mainly by the internal pressure, operating inwardly as a compressive tendency. Ordinarily, osmotic pressures, even when high, are small compared with the internal pressures which counterbalance them. At a semipermeable membrane the osmotic pressure can exhibit itself because it is effective against the membrane, the solute molecules beating a bombardment against the partition through which they cannot pass and which, therefore, operates against the solute as an impermeable wall. On the other hand, owing to the permeability of the membrane for the solvent, the solvent molecules pass freely, and the surface of the solvent, at

which alone internal pressure can exhibit itself, is eliminated at the membrane.

It has been mentioned that the energy effects on condensation of a vapor cannot be explained on the assumption of slowing down the molecules. Indeed, the kinetic theory demands that the effective translational energy of a molecule in the liquid be identical with that in the vapor and a function of the absolute temperature only, *i.e.*, identical for all liquids and vapors at any given temperature level. This important inference can be confirmed by analysis of the nature of osmotic pressure, which can be looked upon as a telltale or device for measuring the effective translational energy of liquid molecules by the behavior of the solute molecules. The gas laws are mathematically equivalent to the assumption that all gas molecules possess an average translational energy dependent upon temperature only and independent of molecular weight and structure. That the osmotic pressure follows the gas laws as long as concentrations are comparable with those at which gases themselves conform points to the conclusion that all solute molecules likewise have effective average translational energies identical with those of gases at the same temperature, and are also uninfluenced by molecular weight and structure. Since this is true of solutes, however varied in type, and since the solute molecules acquire their translational energy only through energy interchange with the solvent molecules around them, there is every reason to believe that the effective translational energy of the solvent molecules must also conform to the same rule. This again compels the conclusion, not only that the phenomenon of liquefaction is not due to loss in velocity of the vapor molecules, but that on condensation no change in effective translational energy occurs.

In the diffusion of a solute from a solution into the pure solvent or into a weaker solution, the same mechanism is operative. The surface of the solution is eliminated and its internal pressure, therefore, rendered inoperative. As a consequence, solute molecules are free to travel over the boundary into the pure solvent, driven by the equivalent of a gas pressure. Osmotic pressure and diffusion are thus interrelated phenomena, having as their basis the kinetic movement of the solute molecules.

Volatility.

The mechanism of liquefaction here discussed gives helpful insight into the effect of molecular weight on volatility. Consider two molecules, one of high molecular weight and one of low. Since the evidence indicates that differences in molecular density, as, for example, in organic liquids, are relatively slight, the large molecule will have a large volume and a correspondingly larger surface. Granting, for the moment, approximately equal

intensities of the stray fields of force around the two molecules, the larger molecule will exert the larger total attractive force because of its larger area.   If it be conceived as resting at the interface between vapor and liquid, it will be attracted toward the liquid more strongly.   On the other hand, its tendency to leave the liquid is due to its translational energy, and this is the same for both molecules.   Consequently, the net tendency for the larger molecule to escape from the surface of the liquid into the vapor will be less than that for the smaller, resulting in a lower vapor pressure.   In other words, at constant temperature the vapor pressure of a liquid should be less the greater the molecular weight, and at constant pressure the boiling point should be correspondingly higher.   Hence, an important factor in determining boiling point is molecular weight; this is well exemplified by the boiling points in any homologous series in organic chemistry.   Thus ethane, with a molecular weight of 30, boils at $-88.3°$ C., while decane, with a molecular weight of 142, boils at 174° C.

However, even superficial familiarity with the facts convinces one that molecular weight cannot be the sole factor in determining boiling point.   Methane, with a molecular weight of 16, boils at $-161°$ C.; ammonia, molecular weight 17, boils at $-33°$; whereas water, molecular weight 18, boils at 100°. The obvious assumption to explain these discrepancies is that there is marked variation in the intensity of the stray fields of force around the molecules.   This factor is more generally known as the polarity of the liquid.   In other words, non-polar methane, with weak fields of force, has a low boiling point, whereas the highly polar water, with intense fields, boils at an absolute temperature over three times as great.   That this picture is correct is confirmed by the parallelism, in a liquid of given molecular weight, between the boiling point and the dielectric constant, which is a reasonably direct measure of the intensity of the electrical forces around the molecule.   Molecular complexity also plays a part, as shown by the fact that certain groups or radicals, such as the hydroxyl group, are always characterized by high polarity.   The volatility of a liquid depends mainly on the combined effects of molecular weight and polarity.

Polarity.

Like internal pressure, polarity does not lend itself to direct measurement or definite evaluation. Liquids, for the sake of convenience, are generally divided into two classes, polar and non-polar, although the boundary line is far from distinct. Polar liquids, as explained above, are those which contain molecules having high mutual attractions or large unsatisfied fields of force.

The dielectric constant * of a compound is one of the best measures of polarity. Polar molecules contain powerful fields of force and therefore tend to orient themselves in an electric field. This greatly increases the capacity of the condenser of which the material forms the dielectric, and consequently the dielectric constant of the material. In Table II are listed the dielectric constants of a number of liquids. By a consideration of data such as these, it has been found that certain groups, e.g., $-OH$, $-NO_2$, $-COOH$, $-NH_2$, etc., impart polarity to a molecule and endow it with an exceptionally strong field of force. If, however, the fields of force can be neutralized within the molecule itself, the forces of intermolecular attraction may be quite small in spite of the presence of polar groups. For example, unsymmetrical nitromethane has a high dielectric constant, while symmetrical tetranitromethane is nearly non-polar.

In some highly polar compounds there are permanently localized electric fields at different points upon the surface. Thus the molecule acts to a greater or less degree as a charged body, having a positive charge at one point and a negative at another. The magnitudes of these two charges, while opposite in sign, will, of course, be equal. The molecule is electrically similar to a short bar magnet and is called a "dipole." Like a magnet, it will develop a moment, which can be measured by taking advantage of the molecule's tendency to turn in an electric field. The possession of a dipole moment by a compound is thus another criterion of polarity, and values are listed with the dielectric constants in Table II. It will be noticed that

* Defined as the ratio of the capacity of a condenser containing the substance in question to its capacity at the same potential when the space between the condenser plates is evacuated.

some substances which possess a low dielectric constant (E) still have a dipole moment (μ). Discretion must always be used in characterizing a liquid as polar or non-polar, but probably the safest criterion is the relation between boiling point and molecular complexity, the latter perhaps best gauged by atoms per molecule.

TABLE II. MOLECULAR WEIGHT AND POLARITY

Substance	Formula	Mol. Wt.	B. Pt., °K.	E (20° C.)	$\mu \times 10^{18}$ e.s.u.
Water	H_2O	18	373	81	1.8
Hydrogen cyanide	HCN	27	299	115	2.6
Ethyl alcohol	C_2H_5OH	46	352	26	1.7
Nitromethane	CH_3NO_2	61	375	39	3.8
Sulfur dioxide	SO_2	64	263	12.8	1.6
Benzene	C_6H_6	78	353	2.3	0.0
Hexane	C_6H_{14}	86	342	1.9	0.0
Chlorobenzene	C_6H_5Cl	112.5	405	5.9	1.5
Chloroform	$CHCl_3$	119.5	334	5.1	1.0
n-Octyl alcohol	$C_8H_{17}OH$	130	467	3.4	1.6
o-Dichlorbenzene	$C_6H_4Cl_2$	147	453	7.5	2.2
m-Dichlorbenzene	$C_6H_4Cl_2$	147	446	8.8	1.5
p-Dichlorbenzene	$C_6H_4Cl_2$	147	446	2.9	0.0
Carbon tetrachloride	CCl_4	154	350	2.2	0.0
Ethyl iodide	C_2H_5I	156	345	7.4	1.7
Bromine	Br_2	160	332	3.2	0.0
Tetranitromethane	$C(NO_2)_4$	196	399	2.1	0.0

Solubility.

In a mixture of two liquids, A and B, in which the force of intermolecular attraction of each species for itself is the same, the force of attraction between the dissimilar molecules A and B will usually be of the same magnitude, granting the fields of force around the molecules are sufficiently symmetrical or non-polar. An example is the mixture of the non-polar liquids, benzene and chloroform, which, since their internal pressures are nearly equal, may be assumed to have nearly the same intermolecular attractive forces. In the contrary case in which the force of attraction between the molecules of B is greater than that between those of A, the molecules of B will tend to associate by themselves and squeeze or elbow out the molecular species A. Hence the two liquids will tend to be incompletely

miscible, as in the case of phenol and water.* The attraction between the molecules of A and B may be greater than that between molecules of their own species, generally due to a highly polar character of the molecules, frequently accompanied 'by convincing evidence of chemical action.   A case in point is that of sulfur trioxide and water.   While attempts to develop a general, quantitative theory of mutual solubility of liquids have met with only limited success, the following qualitative rules are helpful:

In the absence of chemical interaction, liquids of approximately equal internal pressures tend to be miscible and their mixtures show normal vapor pressure relations.

Non-polar liquids of low internal pressure tend to be immiscible with polar liquids of high internal pressure, and, within the limits of miscibility, show abnormally high vapor pressures tending toward formation of azeotropic mixtures of minimum boiling point.

BIBLIOGRAPHY

BRIDGMAN, P. W., *The Physics of High Pressure*, Macmillan (1931).

BRISCOE, H. T., *The Structure and Properties of Matter*, McGraw-Hill (1935).

EDSER, E., "Molecular Attraction and the Physical Properties of Liquids," *Brit. Assoc. Adv. Sci., 4th Report on Colloid Chemistry*, p. 40 (1922).

HILDEBRAND, J. H., *Solubility of Non-Electrolytes*, Ed. 2, Reinhold (1936).

LOEB, L. B., *The Kinetic Theory of Gases*, Ed. 2, McGraw-Hill (1934).

SMYTH, C. P., *Dielectric Constant and Molecular Structure*, Chem. Cat. Co. (1931).

STILLWELL, C. W., *Crystal Chemistry*, McGraw-Hill (1938).

STUART, H. A., *Molekülstruktur*, Springer, Berlin (1934).

* A mixture of ethyl alcohol and water is a case in which the tendency has not developed to the point of separation of phases.

Chapter II

Viscosity

Careful observation of the motion of fluids shows that flow may take place in two ways. If velocity of movement is sufficiently low, the particles of the fluid move in parallel elements, the velocity of each element differing but infinitesimally from that of the adjoining ones. This type of flow is called laminar, strea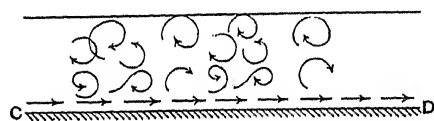mline, or viscous. At sufficiently high velocities the character of the motion changes completely. The orderly, laminar arrangement of streamline flow is replaced by a chaos of swirling eddies and vortices, known as turbulent flow. For the case of a liquid flowing on a horizontal plate CD, Figs. 1(a) and 1(b) represent diagrammatically the velocity distribution in the two types, the arrows as vectors indicating direction and velocity of the fluid elements.
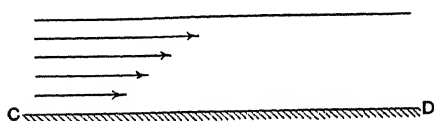
FIG. 1(a).

FIG. 1(b).

If flow is streamline, it is found that in order to maintain the motion of a given element of the liquid at a constant velocity there must be exerted over an area A a shearing stress F, proportional to the velocity gradient* at right angles to flow. In Fig. 2, B and C represent two planes, each of area A, in the fluid, parallel to the direction of flow and x units of distance apart, the arrows indicating rela-

$$\frac{dv}{dx} = \frac{v_1 - v_2}{x}$$
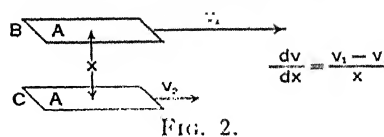
FIG. 2.

---

* The velocity gradient is defined as the rate of change of velocity with distance and is expressed mathematically as dv/dx.

tive velocities. The force F, causing flow, acting on the upper plane relative to the lower is proportional to the area and to the velocity gradient dv/dx, between the planes. Mathematically expressed,

$$F = \mu \, A \frac{dv}{dx}. \tag{1}$$

The constant of proportionality μ in (1) is called the coefficient of viscosity, or simply the viscosity, and is determined experimentally by measuring the flow produced by definite shearing stress under conditions such that the flow is streamline.* The *poise* is the unit of μ when the other quantities are expressed in c.g.s. units. One hundredth of this, the *centipoise*, is often used.

The viscosity is one of the most significant properties of a fluid. The purpose of this chapter is to indicate methods of measurement, to gain an insight into the causes of fluid viscosity, and to discuss the relationships between it and other characteristics of fluids.

The Capillary Viscosimeter.

If a fluid be allowed to flow through a tube of diameter small enough so that the motion is viscous, the flow follows the equation,

$$\mu = \frac{\pi \, \Delta P \, R^4}{8 \, L \, Q} \tag{2}$$

where μ is the viscosity, ΔP the pressure difference at the two ends of the tube, L the length of the tube, R the radius, and Q the rate of liquid flow, volume per unit time † (Poiseuille's

---

* It is important to emphasize that (1) applies only when the motion of the fluid is streamline. However, it must not be inferred that the viscosity of the fluid is without significance in turbulent motion. The viscosity is a fundamental property of the fluid itself. When the motion becomes turbulent other factors are predominant in determining the magnitude of fluid friction, so that other methods of computation are employed, but the viscosity still plays its part and must be allowed for, particularly in estimating the coefficients of friction under such conditions.

† Consider a circular tube of radius R and length L (Fig. 3). Let a fluid be maintained in uniform viscous flow through L by

FIG. 3.

means of a pressure drop ΔP. Consider a portion of the liquid, bounded by a circular tube of length L, of internal radius r, and external radius $(r + dr)$. The force F, effec-
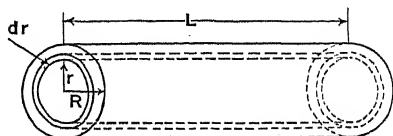
Law). From the equation it is obvious that, if the flow be determined at several different pressures and the results plotted with rate of flow as ordinates against pressure as abscissae, the relationship will be a straight line through the origin, the slope of which is inversely proportional to the viscosity. Since the quantities on the right are all subject to direct measurement, the method is an absolute one.* It is by far the most convenient and satisfactory method to use for liquids and solutions of ordinary viscosities.†

Several forms of the capillary viscosimeter are in use. For work of the highest precision a long, narrow tube is employed, through which a definite volume of liquid flows from one bulb to another under a controlled pressure gradient in a measured time. For work requiring no great precision, modifications of the so-called Ostwald viscosimeter (Fig. 4) are convenient. This consists of a capillary tube inserted between an upper and a lower reservoir. The liquid placed in the upper reservoir flows through the tube under gravity and the time required for a definite volume of flow (fall between marks A and B) is measured. For a given instrument, the time for different liquids is proportional to their viscosities, corrected for differences in den-

FIG. 4. Ostwald Viscosity Pipette.

---

tive in overcoming the viscous resistance of the liquid, will be $F = \pi r^2 \Delta P$, acting over the external surface of the tube, $A = 2\pi r L$.

Hence, (1) becomes

$$- F = - \pi r^2 \Delta P = \mu(2\pi r L)\frac{dv}{dr}. \tag{3}$$

Separation of variables and integration over the limits of r from r to R and v from V to 0, since the velocity of the fluid at the solid wall is zero, gives

$$V = \frac{\Delta P}{4\mu L}(R^2 - r^2). \tag{4}$$

Equation (4) shows that the velocity distribution is parabolic, increasing from zero at the wall of the tube to a maximum value $V_{max.} = \Delta P R^2/4 \mu L$ at the center. However, while it is difficult to determine experimentally the variables of (4), the rate of flow is easily calculated by integration of the variable velocity over the cross section of the tube, giving Poiseuille's law,

$$Q = \frac{\pi \Delta P R^4}{8 L \mu}. \tag{5}$$

* One serious difficulty is securing a cross section exactly circular and uniform.

† For precise work, certain corrections must be applied to take account of energy losses due to effects other than viscosity, particularly losses at entrance and exit of the tube caused by the sudden changes in velocity.

sity.   Hence, the instrument is calibrated with a liquid of known viscosity μ_2 and density ρ_2, and from the times of flow t_1, t_{2}, respectively, knowing the density of the second liquid ρ_1, its viscosity can be calculated:

$$\frac{\mu_1}{\mu_2} = \frac{t_1\,\rho_1}{t_2\,\rho_2}.\qquad(6)$$

Since it is impossible to vary the effective pressure producing flow, only one point on the flow-pressure diagram is obtained. Hence, for liquids of low viscosities it is difficult to make sure that viscous flow is maintained.*

The Rotating Cylinder Viscosimeter.

In dealing with many fluids, particularly those of high viscosity, a convenient form of viscosimeter consists of two concentrically disposed cylinders with the annulus between them filled with the fluid under investigation.   One of the cylinders is rotated at a constant angular velocity and the torque necessary to maintain this rotation is measured.   If the motion of the fluid is viscous, the following relationship holds:

$$\mu = \frac{T}{4\pi\,L\Omega}\left[\frac{1}{R_1^{\,2}} - \frac{1}{R_2^{\,2}}\right]\qquad(7)$$

where Ω is the angular velocity of rotation, T the torque, L the length of the cylinders, R_1 the radius of the inner, and R_2 that of the outer cylinder.   Hence, if the value of Ω is plotted against the torque, there results a straight line whose slope is inversely proportional to the viscosity, the proportionality constant being a function of the dimensions of the instrument.†

FIG. 5

* For routine testing of certain types of liquids, notably lubricating oils, the so-called orifice viscosimeters have been developed, as typified by the Saybolt instrument.   In these a definite volume of liquid is allowed to flow through a short capillary tube.   The time of efflux of the fluid is a measure of the viscosity, but these instruments, while having a restricted usefulness, are not recommended and should not be used for other than the types of liquids for which they are designed, since they aggravate and overemphasize the non-viscous energy losses in flow.

† Consider two long concentric cylinders (Fig. 5) of length L, the inner cylinder having a radius R_1, and the outer R_2. Let the annulus between them be filled with fluid of viscosity μ, and let the cylinders be rotating relative to each other with a uniform angular velocity Ω, so that the fluid between them is in viscous motion.   In order to maintain the rotation a torque T must be exerted on one of the cylinders.   Consider a differential cylindrical shell of
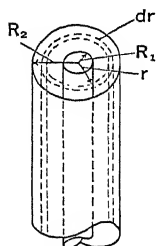
In practice, two instruments embodying this principle have come into wide use, especially in measuring the viscosity of suspensions. The MacMichael viscosimeter employs an outer cup which is rotated at a definite angular velocity, variable at will over fairly wide limits. The inner cylinder is suspended on a torsion wire of known rigidity and the torque measured by the twist. The other instrument is the Stormer viscosimeter, consisting of a stationary outer cup, and an inner cup rotated by a constant torque, impressed by means of a falling weight operating through a cord over a pulley. The angular velocity of the inner cylinder is determined by a stop watch. Instruments of this type, despite lower accuracy, are particularly convenient for routine work, since they are easily cleaned and are capable of giving results over wide ranges of flow conditions with a minimum of experimental detail. End effects can introduce considerable error but are readily allowed for by calibration of the instrument with liquids of known viscosity, determined by other methods. The cylindrical viscosimeter has the advantage of submitting the fluid under test to a nearly constant rate of shear, important in testing materials whose viscosities are a function of the rate of shear (see pp. 167, 220, 224, 243–246).

Falling Ball Viscosimeter.

When dealing with very viscous, transparent fluids, the viscosity can often be determined by measuring the velocity of fall of a sphere through the material. After reaching uniform velocity of fall in viscous motion through a fluid, the following relation holds:

$$\mu = \frac{2}{9} \frac{(\rho_1 - \rho_2)gR^2}{v} \tag{10}$$

the liquid, concentric with the two cylinders, whose inner radius is r, $R_1 < r < R_2$, and whose outer radius is $(r + dr)$.

The force causing relative motion, acting on the inner surface of this shell, is $F = T/r$. At this point the velocity gradient is $dv/dr = r \, d\omega/dr$. Hence, (1) becomes

$$\frac{T}{r} = \mu(2\pi r \, L)r\frac{d\omega}{dr} \tag{8}$$

where ω is the angular velocity of the fluid itself. Separation of variables and integration between the limits $r = R_1$, $\omega = 0$ and $r = R_2$, $\omega = \Omega$ give

$$\Omega = \frac{T}{4\pi \mu L}\left[\frac{1}{R_1^2} - \frac{1}{R_2^2}\right]. \tag{9}$$

where ρ_1, ρ_2 are densities of sphere and fluid, respectively, R the radius of the sphere, g the gravitational constant, and v the velocity of fall.* In using this method the dimensions of the containing vessel should be large compared with those of the sphere, although one can use the corrected formula †

$$\mu = \frac{2(\rho_1 - \rho_2)gR^2}{9v\left(1 + 2.4\dfrac{R}{R_1} + 2.4\dfrac{R^2}{R_1^2}\right)} \tag{11}$$

where R_1 is the radius of the containing vessel. For liquids of ordinary viscosities it is difficult to obtain accurately dimensioned spheres of small enough size. In a recently developed viscosimeter this is overcome by allowing a ball or sphere to roll down through the liquid in a closely fitting glass tube, placed at an angle to the horizontal. The velocity of the sphere is a complex function of the viscosity, which is determined by calibration with liquids of known characteristics. This instrument, like the Ostwald viscosimeter, suffers from the serious disadvantage of determining only one point on the rate-of-flow *vs.* pressure-difference curve.

TABLE I. VISCOSITY OF FLUIDS AT 20° C. IN CENTIPOISES ‡

Acetic acid	1.23	Heptyl alcohol	7.01
Air	0.0181	Hexane	0.320
Ammonia	0.0108	Hydrogen	0.0088
Benzene	0.647	Methyl alcohol	0.593
Carbon dioxide	0.0160	Nitrogen	0.0184
Glycerol	1069	Oxygen	0.0206
Aniline	4.43	Valeric acid	2.30
Chloroform	0.563	Water	1.01

A liquid sufficiently viscous can give the impression of solidity and yet flow (pp. 297–299). Thus, a rod of such a material of length L, under the tension of a total axial force F, the stress uniformly distributed over its cross section A, elongates at a rate $dL/dt = FL/3\mu A$.

---

* For derivation of this formula, see p. 98.
† BARR, GUY, *A Monograph of Viscometry*, Oxford University Press, London (1931), p. 182.
‡ Data from *International Critical Tables*, Vols. V and VII, McGraw-Hill.

Consider * a point in a circular rod of radius R at axial distance z from one end, radial distance r from the axis and angle θ from a fixed reference direction. Call the axial component of the velocity $w = c_1 z$ and the radial component, $u = c_2 r$. The component stresses are σ_z, σ_r, and σ_θ. The relations are

$$p - 2\mu \frac{\partial w}{\partial z}; \qquad \sigma_r = p - 2\mu \frac{\partial u}{\partial r}; \qquad \sigma_\theta = p - 2\mu \frac{u}{r}.$$

Replacing the partial differential coefficients of the second group of equations by c_1 and c_2 by substitution from the first group, followed by elimination of p, gives

$$\sigma_z = 2\mu(c_2 - c_1) \cdot \frac{F}{\pi R^2}.$$

The condition of incompressibility is

$$\frac{\partial w}{\partial z} + \frac{1}{r}\frac{\partial}{\partial r}(ru) = 0,$$

whence $c_1 + 2c_2 = 0$ and $\sigma_z = -3\mu(\partial w/\partial z)$. Since $dL/dt = w_{z=L}$,

$$\frac{dL}{dt} = \frac{FL}{3\mu\pi R^2} = \frac{FL}{3\mu A}.$$

This relation has been used to measure high viscosities by measuring the rate of change of distance between two marks in the middle † of a thread supporting a known weight (see p. 295).

The Significance of Viscosity.

The viscosities of the two types of fluids, liquids and gases, exhibit outstanding differences (Table I):

1. Isothermal changes in density of a gas at low pressures have practically no effect on the viscosity, whereas the influence of the density of a liquid on its viscosity is great.

2. The viscosity of a gas is always small, whereas that of a liquid is relatively high.

3. At a given level of temperature the differences in viscosity of various gases are relatively small—a few fold at the most—while the differences between liquids are enormous. On the other hand, the most diverse pure liquids differ in viscosity no more than the gases if their viscosities are compared at their normal boiling point temperatures.‡

* The authors are indebted to Dr. E. Reissner, Department of Mathematics, Massachusetts Institute of Technology, for this derivation.

† This avoids end effects caused by unequal distribution of stress.

‡ It is sometimes stated that kinematic viscosities (ratio of viscosity to density) are more nearly equal at atmospheric boiling points, but it is doubtful if the comparison is better than between the viscosities themselves.

4. The viscosity of a gas increases with temperature, whereas that of a liquid decreases. This is true even when the change in temperature is made at constant density, by adjustment of the pressure.

Figure 6 represents a gas flowing in streamline motion, the arrows indicating the velocity of flow of neighboring elements. In each element the molecules are moving in the irregular disordered motion of thermal translation, but superimposed upon this are the directed components of velocity indicated in the figure. Owing to the thermal agitation, molecules are continually passing from one element to another. As a molecule, such as G, moves from a zone of high velocity CD into one of low EF, it collides with molecules in the latter. While a part of its former directed energy is imparted to the molecules of the more slowly moving element, a part is dissipated by conversion to *undirected* energy, not only of molecular translation, but also of molecular rotation and vibration. Similarly, if the molecule G wanders into the high velocity stream AB, there is a net slowing down of the molecules in the latter and a speeding up of G, but this again results in a net dissipa-

FIG. 6.   FIG. 7.

tion of directed into undirected energy: *i.e.*, a conversion of work into heat. This loss of directed energy is the friction of viscous flow.
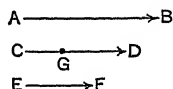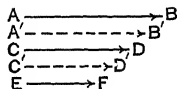
Assume that the gas in Fig. 6 has its density doubled at constant temperature, under conditions maintaining the original velocity gradient, as in Fig. 7. This results in doubling the number of molecular collisions and correspondingly tends to increase the dissipation of directed energy. However, the molecules have, on the average, only half as far to move between collisions, and therefore possess on collision only half the difference in directed velocity that they had before the density was increased. Since the velocity gradient is unchanged, the energy dissipation per impact is correspondingly lessened. It is clear that the effects of the increased number of collisions and the decreased differences in directed velocity between elements tend to cancel. Analysis of the mechanism of energy loss shows that this cancellation is quantitative so long as a perfect gas is con-

sidered, its viscosity being independent of density and therefore
of pressure.   Similarly, increase in gas temperature increases
frequency of collision, and hence increases viscosity.

However, there are factors, insignificant in their effects at
low densities but important at high, which influence the viscosity
characteristics even of gases and become controlling in liquids.
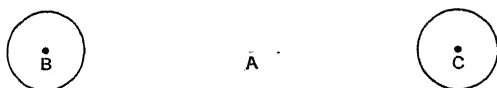In Fig. 8 consider three molecules in a gas, molecule A having

FIG. 8.

just collided with B and rebounded in the direction of C.   Around
B and C is indicated the extent of the sphere of action.   For
the major portion of the travel of A between B and C, A is
outside the influence of the molecular forces of B
and C and is therefore traveling at constant veloc-
ity, namely, the thermal velocity discussed in
Chapter I.   Figure 9, on the other hand, indicates
the situation existing in a liquid.   The spheres of
action intersect, not only when the molecules are in their aver-
age position, but even at points of maximum separation.
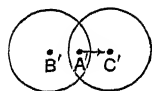
FIG. 9.

Assume that in a gas A has just collided with B and is re-
bounding in the direction of C.   In Fig. 10 the distance between
the two molecules A and C is plotted along the line GH, and the
corresponding mutual attractions are represented by the curve
D and the repulsions by E, neglecting for the moment the effect
of the presence of surrounding molecules.   The net resulting
intermolecular force is given by the dotted curve F.   When, on
approach, the molecules reach the point represented by I they
clearly start to accelerate towards each other, but at K this
acceleration becomes zero; the approaching molecules slow down
rapidly and start to recede.   When the point K is again reached
during recession, the separating molecules have enough kinetic
energy to overcome the attractive forces throughout the zone
KI, so that the molecules pass through the point I with the
effective thermal velocity characteristic of their temperature

level.* The changes in relative velocity, corresponding to Fig. 10, are shown on Fig. 11.†

In a liquid, the molecule A moves through the neutral zone with a translational energy equal to that in a gas, namely, $3/2\ RT$, and hence with a velocity identical with that which it would possess on the average in the gaseous state. However, for most of the time between collisions its velocity, as shown by Fig. 11, is substantially above that in the neutral zone. Hence, in a liquid the average velocity of a molecule is higher than at the same temperature in the gaseous state. Therefore, the number of collisions experienced by a molecule in unit time will be higher in a liquid than in a gas, all the more because of the fact that the distance between collisions (mean free path) is greatly reduced, since the voids between the molecules, through which alone they can move, are only a modest fraction of the total volume. This large increase in the frequency of collisions in the liquid induces a corresponding increase in the dissipation into heat of the directed energy of laminar motion, contributing to the difference in gaseous and liquid viscosities.
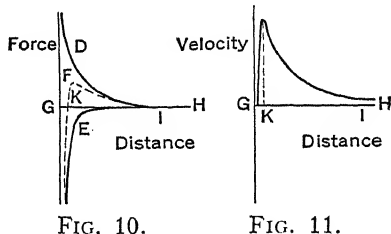
FIG. 10.　　　FIG. 11.

In a sense this is equivalent to assuming that the major movement of the molecule in a liquid is vibratory, bouncing back and forth rapidly over a path that changes relatively slowly. The structure is thus analogous to that of a solid in which atom or molecule, as the case may be, must, normally at least, vibrate about a fixed mean position. However, it is essential to keep in mind the fact that in the liquid the molecule is free to wander and is continually changing its position. Furthermore, the energy per molecule available to cause displacement is identical with that in a gas at the same temperature, as evidenced by the conformity of osmotic pressures to the gas laws. The rate of displacement of the individual molecule in a liquid (e.g., rate of diffusion) is less than in a gas because of greater fric-

* The presence of neighboring molecules, as in a liquid, complicates this picture somewhat, but does not change the general shape of the curves.

† The translational energy of A, which is available for any external effect (such as osmotic pressure in a liquid), is only that corresponding to its kinetic energy of translation as it passes through the neutral surface, corresponding to the point I of Fig. 10, since its other energy is bound and unavailable for collision purposes with external objects (e.g., for osmotic expansion work at a semipermeable membrane).

tion, not lower driving force. In a liquid there is apparently no static friction of continuing molecular displacement, however high the friction may become at finite displacement velocities. In consequence, it is to be expected that the arrangement of the molecules in a liquid is largely, though perhaps never wholly, random (see pp. 143–144).

The superimposed effect on frequency of collision of the increase in average molecular translational velocity and the decrease in void volume around the molecules which accompany liquefaction of a vapor must be very great. The fact that the viscosity of the liquid at its boiling point is only some 30-fold that of its vapor is therefore surprising and indicates that the dissipation of directed energy of movement resulting from each collision of a molecule from a slow-moving filament of flow with one from its neighboring, faster moving lamina is less in the liquid than in the gas. The relative structures of the two fluids offer explanation. In the case of a gas, almost all of the molecules moving from the low velocity filament into the neighboring high one will penetrate it before collision occurs, get into the line of fire and, on the average, be overtaken from behind by a faster moving molecule. This results in a substantial equalization of their directed velocities. In a liquid, on the other hand, because the molecules are closely packed together, the collision of molecules from neighboring filaments of flow is usually sidewise, *i.e.*, along a line of centers nearly at right angles to direction of flow. If collision is exactly at right angles, rebound is rapid and there is no slipping friction between molecules in contact; there will be no interchange of directed energy of motion and hence no development of fluid friction. The impossibility of full penetration of the molecule from one lamina of flow into another greatly lessens the effectiveness of collision in inducing viscous friction.

However, the correlation of liquid viscosity with polarity and molecular weight (see p. 16) indicates that slipping friction between contacting molecules is often high. Molecules with strong, interlocking fields of force apparently can "stick" sufficiently on even sidewise impact so that one molecule can drag the other along in the sense of imparting to it some of its kinetic energy of translation in a direction parallel to the surface of

contact, even though rebound in a direction perpendicular to that surface quickly separates them. In so far as these mutual attractions between the molecules of the liquid are counterbalanced by the disgregating effect of thermal agitation, they lose their capacity to induce high viscosity. This is why at the boiling point, where these attractions are sufficiently counterbalanced to enable the liquid molecules to escape into the vapor against atmospheric pressure, the viscosities of the most diverse liquids are also reduced to nearly equal values. The correlation between viscosity and vapor pressure is substantially the same whether high boiling point is due primarily to high molecular weight (*i.e.*, large molecular surface for attraction with no increase in thermal disgregation energy per molecule), or to high polarity (*i.e.*, strong attraction per unit surface) in a small molecule. The parallelism is further emphasized by the similarity in the approximation formulae often used in the two cases,

$$\frac{d \ln p}{dT} = \frac{A}{T^2}$$

for vapor pressure, and

$$\frac{d \ln \mu}{dT} = -\frac{\alpha}{T^2}$$

for viscosity.

Temperature Coefficient of Viscosity.

If one can assume the formula of the preceding paragraph, it is clear that a plot of the

FIG. 12. Temperature-Viscosity Curves for Various Liquids.

logarithm of the viscosity of any given liquid against the reciprocal of the corresponding absolute temperature should give a straight line. The data for five typical liquids are thus plotted in Fig. 12.

Inspection shows that there is some curvature, but, neglecting this for the moment, there is a marked trend in the sense that the curves for different liquids get steeper as the viscosity gets higher. This can be seen even more clearly from Table II, the last column of which gives the slopes of such curves.* The study of large numbers of liquids demonstrates that this tendency, while not quantitative,† is in no sense exceptional. The temperature coefficient of viscosity of a liquid evidently depends not only on the temperature level but also on the viscosity itself. If one plots for various organic liquids at normal levels of temperature the fractional change in viscosity per degree of temperature rise, $d\mu/\mu\,dt = d\ln\mu/dt$, against the viscosity on a logarithmic scale, while there is considerable scattering of the points, they are well represented by the curve of Fig. 13.‡

Fig. 13. Change in Viscosity (Poises) with Rising Temperature.

Figure 13 was constructed from data on liquids at normal temperature levels. Few data are available for pure liquids at other temperatures but

TABLE II

Substance	Temp., °C.	Viscosity, in Centipoises	$d\mu/dt$	$\dfrac{d\ln\mu}{dT}$ (exptl.)	$\dfrac{d\ln\mu}{dT}$ (Fig. 13)	$\dfrac{d\ln\mu}{d(1/T)}$
Octane	25	0.51	0.006	0.012	0.015	1,060
Carbon tetrachloride	25	0.90	0.014	0.015	0.018	1,290
Water	25	0.91	0.022	0.024	0.018	2,100
Butyl alcohol	25	2.55	0.064	0.025	0.025	2,200
Phenol	25	9.0	0.41	0.046	0.036	4,100
Glycerin	20	880	78	0.089	0.085	7,600
Glucose	67	2,700,000	510,000	0.19	0.17	22,000

*It is well to note that for these liquids the temperature coefficient of viscosity itself varies 85 million-fold. In this table signs of the slopes are ignored.

† The crossing of the curves for water and carbon tetrachloride on Fig. 12 is, of course, itself an exception.

‡ Were the relation of Fig. 13 exact rather than approximate, it could be integrated into a universal viscosity curve, as in Fig. 14. This latter plot does in fact give a preliminary estimate of the magnitude of change in viscosity (at ordinary temperature levels) one may expect in an organic liquid whose viscosity is known at one temperature.

these indicate that as temperature changes the correlation of $d \ln \mu/dT$ with μ is affected, in that for a given value of μ the differential coefficient is roughly inversely proportional to the 1.5 power of the absolute temperature. This is equivalent to dividing the abscissae of Fig. 13 by $0.0002T^{1.5}$ before using them, T being in degrees Kelvin. To indicate the order of dependability of Fig. 13, the values of $d \ln \mu/dT$ read from it are inserted in the sixth column of Table II, for comparison with the experimentally determined values in the preceding column.

There are few data on pure liquids of high viscosity. As pointed out, viscosity is always low at the boiling point. As temperature is decreased below this level, most liquids freeze before the viscosity becomes very great.

FIG. 14. Viscosity-Temperature Relation of Liquids.

Figure 13 is in no sense universal. Mercury is completely out of line with it. Many liquids, such as water, are abnormal, probably due to extensive molecular association in the liquid state. Molecular shape is a factor of undoubted significance (pp. 164–169) for which it does not allow. None the less, it indicates significant trends and is often helpful in foreseeing them in cases of inadequate data.

Granting that the high viscosity of a liquid is caused by the internal pressure which results from the mutual interattractions of the molecules, the reason for the high negative temperature coefficient of viscosity and its correlation with viscosity itself is easy to understand. That the increase in thermal disgregating pressure of a liquid resulting from a given rise in temperature is very great is indicated by the high increase in pressure on heating at constant volume $(\partial P/\partial t)_v$, since the change in the otherwise counterbalancing internal pressure is not likely to be great under these conditions because average intermolecular distance remains the same. When the liquid is heated without the large increase in external pressure required to keep it at constant volume, its thermal expansion, though relatively small, reduces the internal pressure greatly, because of the influence of molecular distance on the attractive forces (p. 10). The large decrease in viscosity is caused by this drop in internal pressure. The higher the internal pressure, and hence the viscosity, the more rapid the drop with temperature. However, since the increase

in molecular energy per degree rise in temperature available for
thermal disgregation effect is independent of temperature level,
but the corresponding total energy is proportional to the tem-
perature, the percentage increase per degree of thermal disgre-
gating pressure must be less at high temperatures than at low.
Hence, at high temperatures the viscosity reduction per degree
for a given viscosity level will be less.

However, that the relationships are still more complicated than here
indicated is confirmed by the fact that the viscosity of a liquid heated at
constant volume goes down, though far less than if heated at constant
pressure. At constant volume the intermolecular void volume cannot
change enough to make a significant change in the amplitude of molecular
travel between actual collisions, yet the velocity of molecular movement
must increase with temperature. Since this must result in greater fre-
quency of collision and rebound at higher temperature, lowered viscosity
must mean less degradation of energy of directed movement to heat per
collision. There is evidence that the magnitude of the intermolecular
attractions may drop somewhat as temperature rises, but it seems difficult
to explain the often large viscosity effects under discussion on this basis
alone. Greater rapidity of rebound after molecular collisions at high tem-
perature may leave less chance for transfer of directed momentum.

At very high pressures, abnormal viscosity effects are often encountered
which may be due to close-packing and even distortion of the molecules
themselves. Extensive molecular association in a liquid can probably
induce similar results. While the general nature of the factors influencing
liquid viscosity are understood, the details are far from clear.

Based on an analysis of these relationships Andrade has
developed the following equation for the effect of temperature
and density on the viscosity of a given liquid:

$$\mu \sqrt[3]{v} = A \ e^{\frac{c}{vT}}. \tag{12}$$

In this equation, v is the specific volume at the absolute tem
perature T, and A and c are constants specific for each liquid
The equation is probably the best available for pure liquids.
It is clear that by its use one can construct the entire viscosity-
temperature curve if one knows the viscosity of the liquid at
two temperatures and the change in liquid density with the
temperature. However, in the case of specific liquids abnor-
malities are sometimes encountered.

While liquid viscosity, depending as it does on intermolecular attractions, is profoundly influenced by polarity and molecular weight, the former will in general be of greater significance. Thus, referring to Table I (p. 23), and comparing hexane (m.w. = 86) with glycerin (m.w. = 92), it is seen that the highly polar glycerin has 3000 times the viscosity of non-polar hexane at 20° C. However, non-polar materials of high molecular weight, such as certain hydrocarbons, may have viscosities of the order of 10^5 to 10^7 and higher, and yet behave as true liquids. As molecular weight increases, molecular surface increases correspondingly. Since, for a given type of molecule, the intensity of the surface forces per unit area is approximately constant, the total mutual attractive forces per molecule increase greatly with molecular size. The aggregating effect of these molecular attractions is counterbalanced primarily by the disgregating tendency of thermal agitation. Since, however, the average effective kinetic energy of translation per molecule is independent of molecular weight, the excess of the attractive aggregating forces over the disgregating effects increases greatly with molecular size. Attention has already been called to this as an explanation of the decrease in volatility with molecular weight. It is likewise the underlying cause of the influence of molecular weight upon viscosity. Because molecular weight can reach such high values, the ultimate effects on viscosity can become very great indeed. When high polarity and high molecular weight are both present, as in amorphous silica, viscosity at normal temperatures reaches astronomical figures (see, however, pp. 155–158, 283–284, 420–422).

BIBLIOGRAPHY

DA C. ANDRADE, E. N., *Phil. Mag.* 17, 497, 698 (1934).

BARR, G., *Monograph of Viscometry*, Oxford University Press, London (1931).

BINGHAM, E. C., "Viscosity," *U. S. Bureau of Standards, Scientific Papers*, No. 278 (1916).

GEMANT, A., Articles, *J. App. Phys.*, 1941–42.

HATSCHEK, E., *Viscosity of Liquids*, D. Van Nostrand (1928).

SCOTT, G. W., *An Introduction to Industrial Rheology*, Churchill, Ltd., London (1938).

Chapter III

Surface Tension

The interface separating a pure liquid and its vapor is ordinarily conceived as a mathematical surface, with sharp, abrupt changes in physical properties across it; however, it is not impossible that transition from one phase to the other is sufficiently gradual so that there is continuity of properties. The optical properties of a liquid surface indicate that the transition is extremely sharp and abrupt. Rayleigh showed that, if a liquid surface be scrupulously cleaned from all contaminations, such as films of grease, etc., light reflected from the surface is completely plane polarized, no ellipticity being detectable by his technique.* From this it may be estimated that the transition from liquid to vapor takes place predominantly in a layer, approximately one molecule in thickness, a result in complete accordance with the concept of intermolecular attraction, since intermolecular forces are found to vary inversely with a high power of the distance.† On account of this extreme density gradient across the boundary layer between a liquid and its vapor, it is not surprising to find that the interface exhibits certain peculiar properties, of which the phenomenon of surface tension is perhaps of greatest importance.

Surface Tension.

The surface tension of a liquid can be studied experimentally with the simple arrangement shown in Fig. 1(a), which, although

---

* When the transition between two phases is absolutely abrupt, Fresnel's Law of Reflection requires that, if a ray of light has an angle of incidence of $\tan^{-1} \eta$ (where η is the refractive index of the liquid), on reflection it will be completely plane polarized. If the transition is gradual, the light will be elliptically polarized.

† However, sharpness of transition should disappear completely at the critical point, where liquid and vapor become identical and distinctions between the phases disappear. As the two phases approach the critical, the surface layer will become diffuse, the vapor approaching the liquid in concentration. The phenomena are confused by the high compressibility of the fluid near the critical point.

unsuited for precise measurements, is highly instructive.   It
consists of a light wire framework $ABCD$ crossed by a similar
light wire EF, free to move over the frame.   By dipping into
a soap solution, a film can be formed over $ABCD$ in the plane
of the frame, and the behavior of EF watched when one side
of the film, say $EBCF$, is broken.   When the film $EBCF$ col-
lapses, EF moves with great rapidity towards AD.   However,
by holding EF by hand while
breaking the film $EBCF$, the
film $AEFD$ may be maintained
or even extended in area.   A
physical force is evidently re-
quired to prevent the film
$AEFD$ from contracting.   This
force may be investigated by
attaching EF to the fixed cross-

FIG. 1.

piece BC, with a piece of light elastic GH, and cotton threads
HK, as shown in Fig. 1(b) and Fig. 2.   If a film be now
formed on $AEFD$, the elastic is stretched and the cross-wire
comes to rest in a definite position.   While with a new elastic
the cross-wire EF may come to rest in another position rela-
tive to AD, the force or pull exerted by the elastic is found
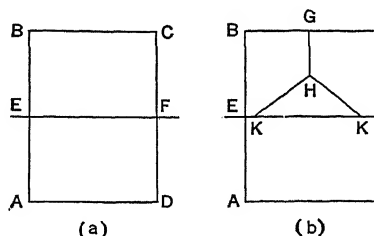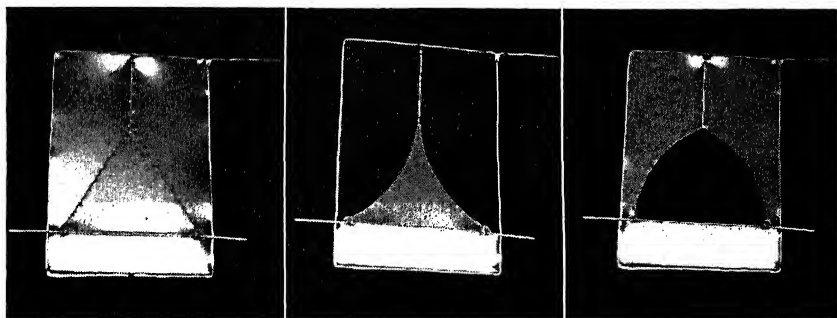
FIG. 2.   Photographs of Soap Films.

the same as in the first case, i.e., the force required to counter-
balance that exerted by the film is independent of any variation
in film area.   By changing the width of the framework, it can
be shown that the force is proportional to the width of the

rectangle AD. Similarly, the force is independent of the thickness of the film. If the area $KKII$ between the cotton threads and the cross-wire is also made to take up a soap film, it is found that the cotton threads are bowed inwards. On breaking the film, the threads straighten and the cross-wire moves to a new position of equilibrium nearer AD, but with the elongation of GII unaffected. It is still more illuminating to place a loose loop of light thread in the film and then rupture the film inside the loop. The thread is immediately jerked into the form of a circle. The preceding facts definitely show that a liquid film is capable of exerting a tension. This tension or force is a purely surface effect, since it is independent of the thickness of the film, and it acts parallel to and in the plane of the surface. It is proportional to the width of the rod and is independent of the distention of the surface. Furthermore, in view of the behavior of the cotton loop, the force must be the same in all directions in the surface. Since the force must be counterbalanced by a pull on the side of the rectangle AD, it must be transmitted through the film as a whole and operate at every point in it. This conclusion is even more convincingly shown by the behavior of curved films, *e.g.*, a soap bubble or a sessile drop, where the shape of the surface always conforms to that required by the presence of such a force at each point in the surface. This force, characteristic of a liquid surface, is called the surface tension and may be defined as the force exerted, in the plane of the surface, per unit length taken perpendicular to the direction of the force * (see below).

Surface Energy.

The behavior of the liquid surface must be considered from another important angle. The methods of measurement just

---

* It should be emphasized that the concept of surface tension, although extremely useful, may occasionally be misleading; thus, surface tension, unlike other tensions, is independent of the amount of distention, *i.e.*, the amount of free surface. The difference between the tension in liquid surfaces and the stresses in solids is clearly realized. Maxwell,[1] for example, states: "The surface must therefore act like a sheet of india-rubber when extended both in length and breadth, that is, it must exert surface tension. The tension of a sheet of india rubber, however, depends on the extent to which it is stretched and may be different in different directions, whereas the tension of the surface of a liquid remains the same however much the film is extended, and the tension at any point is the same in all directions."

[1] Maxwell, J. C., *Encyclopaedia Britannica*, Ed. 9, 5, 66 (1878).

described indicate clearly: first, that isothermal extension of the surface is impossible without the expenditure of mechanical work, and second, that the extension is reversible, the force required being substantially the same for both extension and contraction of the film.   It follows that the work consumed in isothermal extension of the surface is free energy in the thermodynamic sense.   Consequently, isothermal extension of the surface is impossible without the expenditure of energy in the form of work, or an equivalent change in energy availability. Because the energy under discussion is free and therefore recoverable, it must reside in the surface or in its immediate neighborhood.   In magnitude, it must equal the reversible force times the distance through which it operates.   If the surface be enlarged along a line of length L, the force is σL.   Calling the distance moved, measured at right angles to the line of the expanding force, dl, the work done is $dw = \sigma L\, dl$.   Obviously, however, $L\, dl$ is the increase in surface area, $i.e.$, the reversible expansion work is $dw = \sigma\, dA$.   Since under many conditions the surface tension is constant, $w = \sigma A$, $i.e.$, the reversible work required to create the surface is proportional to it.   This reversible work can obviously be expressed per unit area, $i.e.$, $W/A = \sigma$.   This work of surface extension stored per unit area of the film is defined as the free surface energy; it is numerically equal to the surface tension,* but the distinction between the two should always be kept in mind.   From many points of view the free surface energy is the more helpful concept.   It will be remembered that any system tends to assume a condition of minimum free energy.   Thus, in mechanics the principle that static equilibrium is the condition corresponding to a minimum potential energy for the system as a whole is a familiar illustration of the broader generalization.   As applied to the behavior of liquids, it reduces to the principle that a body of liquid always tends to assume the form involving minimum surface compatible with the external restraints to which it is subjected.   The solid figure having the smallest area for a given mass is a sphere: hence, a liquid free from the influence of other forces always tends to assume a spherical shape.   No matter how distorted,

---

* Surface tension can be defined as the intensity factor in the free surface energy.

providing it is not subdivided, the drop always contracts to its original spherical shape on removal of the distorting force.

Total Surface Energy.

The free surface energy does not give the whole of the energy expended when a new surface is formed. If a molecule of a liquid be evaporated from the interior of a liquid through the surface into an evacuated space, work must be done against the attractive forces of the other molecules, *i.e.*, energy is consumed, the total of which is called the internal energy of vaporization.* It can all be supplied isothermally as heat; in other words, it is known thermodynamically as bound energy. A molecule moving from the interior of a liquid to its surface is also being brought from a region of high to one of low density, but the process is somewhat more complicated, although it is easily seen that energy will be required. Under isothermal conditions part of the energy must be supplied in the form of work (or its equivalent) in order to draw apart the molecules in the surface against their mutual attractions, so that the new molecule can find a place, *i.e.*, the free surface energy. However, having prepared a place for the molecule in the surface, energy is still required to bring the molecule to the surface against the molecular attractions in the liquid. This is similar to the case of evaporation, and the energy required can presumably be supplied in the form of heat, *i.e.*, as bound energy. The thermodynamic relation between the total and the free energy is given by the Helmholtz equation,

$$U = \sigma - T \frac{d\sigma}{dT} \qquad (1)$$

where, in this case, σ is the free surface energy, U the total surface energy, and T the absolute temperature. Since $d\sigma/dT$ is negative, it follows that if additional surface is formed adiabatically, heat must be absorbed from the interior of the liquid and its temperature will drop.

A free energy is almost never measured directly, but some intensity factor of the free energy is chosen for measurement,

* This process, unlike the following one, does not require change in the interfacial area.

e.g., pressure, voltage, etc. It is not surprising, therefore, to find that in this case also the quantity measured is not the surface energy itself but the force of surface tension.

The Ring Method for Determining Surface Tension.*

A simple, effective method of measuring surface tension is the torsion balance, perfected by du Noüy, illustrated in Fig. 3.
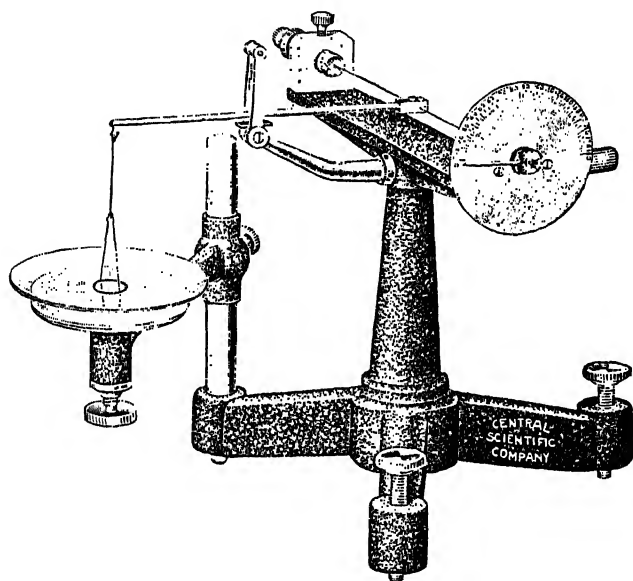
FIG. 3.   du Noüy Tensiometer.†

A horizontal wire, held firmly under tension, has fastened to it at right angles a light arm, also horizontal. The end of this arm carries wires supporting a horizontal ring of thin platinum wire. The ring is dipped in the liquid surface and can be lifted out by twisting the horizontal torsion wire, which is calibrated to measure the necessary force. The rising ring brings with it a film of the liquid. Initially this film tends to be cylindrical in form, and, since it has a double surface, will have a total length equal to

---

* All methods of determining surface tension are extremely sensitive to small amounts of impurities. In particular, the surface tension of many solutions varies with the age of the surface, mode of formation, etc., factors which are important in the study of adsorption and allied problems (pp. 56–66).

† Central Scientific Company, Bull. 101.

twice its circumference. Thus the total force due to surface tension tending to drag the ring back into the liquid will be $4\pi r\sigma$, where r is the radius of the ring and σ the surface tension per unit length of film. The force necessary to oppose this is measured by the angle of twist, and, from a preliminary calibration of the wire, can be read directly in dynes.*

A similar approximate method is the measurement of the pull required to lift a microscope slide out of the surface of a liquid. The slide is suspended edgewise from a balance beam so that its long, narrow edge just touches the liquid. The pull upon it due to surface tension will be equal to $2(a + b)\sigma$, where a and b are the lengths of the side of the edge touching the liquid. Weights are added to the other pan of the balance until the slide is just pulled from the liquid, when the excess weight must be equal to the pull due to surface tension.

These methods, although not of the highest accuracy, are illustrative, and the first is important as a rapid and convenient method of determination.

The Maximum Bubble Pressure Method.

Within a bubble of air suspended in a liquid, as shown in Fig. 4, there must be an excess pressure h, tending to expand it and acting against and counterbalanced by the surface tension at the interface. In any meridian plane AB, the surface tension will act along the circumference with a total force $2\pi r\sigma$, r being the radius. This force must be equal to that due to the pressure difference p, acting over the area of the section; or,
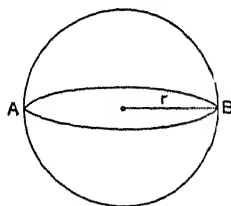
FIG. 4.

$$\pi r^2 p = 2\pi r\sigma \quad \text{or} \quad p = \frac{2\sigma}{r}. \qquad (2)$$

For a soap bubble there will also be an outside film, so that the pressure within the bubble will be double this value. Measurement of the radius of a bubble is comparatively easy, but inser-

* For accurate work a correction, empirically determined, due to distortion of the shape of the film and perhaps other factors, must be applied to this simple formula. [HARKINS, W. D., and JORDAN, H. F., *J. Am. Chem. Soc.* **52**, 1751 (1930); FREUD, B. B., and FREUD, H. Z., *J. Am. Chem. Soc.* **52**, 1772 (1930).]

tion of any pressure-measuring or transmitting device distorts the bubble and allowance must be made therefor.   The bubble can be formed at the end of a capillary tube, as in Fig. 5.   Air is slowly expelled from the tube, forming a bubble, which finally becomes hemispherical.   Up to this point the radius of curvature of the expelled bubble is progressively decreasing, but further growth is impossible without increase in the radius at some point.   This renders the bubble unstable and it immediately breaks off.   At this point the pressure, h_{max}, which is a maximum, is read.

FIG. 5.
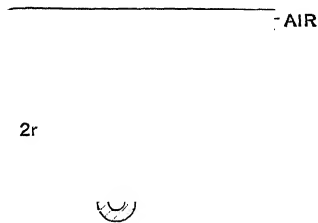
If the capillary tube has an inner radius r and is immersed vertically to a distance a in the liquid of density ρ, then at this moment

$$h_{max} = \frac{2\sigma}{r} + g\rho\,(a + r),\qquad(3)$$

from which the surface tension can be readily calculated.*   Because the method † requires only small quantities of material, simple apparatus, and measurements can be made over a wide temperature range, it is coming into increasing favor.

The Drop-Weight Method.

If a liquid be allowed to form a drop slowly in air (or any immiscible fluid) at the end of a vertical tube, it hangs pendent, gradually forms a constriction just below the point of attachment to the tube, and finally breaks away at this narrow neck. If movement be sufficiently slow, there is equilibrium between the net weight of the drop and its surface tension acting around the line of minimum cross section.   Were the radius of the minimum cross section that of the tube, the relation would be

$$mg = 2\pi r\sigma\qquad(4)$$

* Clearly, r should be the radius of curvature of the bubble tip, but difference between this and the radius of the tube is negligible for small bubbles.   With large tubes this simple formula does not hold, for, due to hydrostatic head effect, the bubble is no longer hemispherical at the moment of rupture.

† SUGDEN, S., *J. Chem. Soc.* **121**, 858 (1922); *ibid.* **125**, 29 (1924).

where m is the mass of the drop, r the radius of the tube, and σ the surface tension of the liquid. In practice it is found that the drops always have a weight varying according to conditions,

FIG. 6.

but, due to constriction below the tip (Fig. 6), usually smaller than $2\pi r\sigma$ by about 40%.
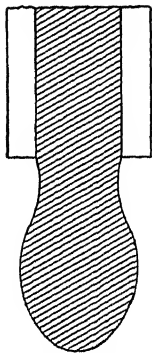
The photographs of a falling drop taken with a high speed motion picture camera and reproduced in Fig. 7 show convincingly why this is the case. The pendent drop does not split off cleanly, but a number of satellite drops are formed, inevitably affecting the energy relationships involved in any calculation and showing that a complicated dynamic analysis of the problem would be required. However, as in other cases, e.g., the flow of fluids in turbulent motion, for which direct mathematical analysis is difficult, the problem can be approached and solved by the use of dimensionless ratios. Harkins and
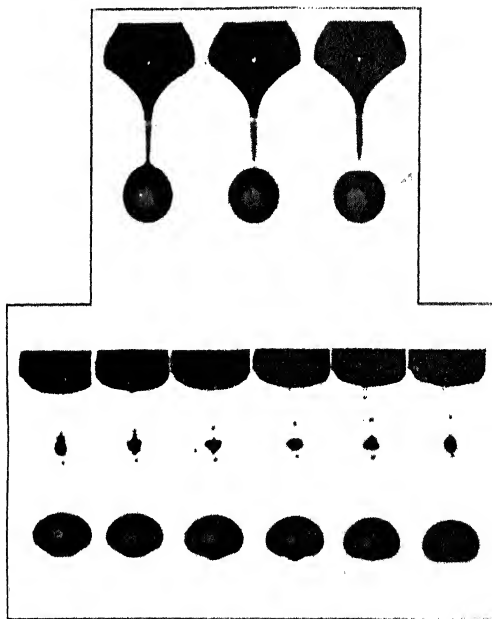
FIG. 7. Drops of Water Falling from a Tip.*

Brown, by a long series of careful experiments, have succeeded in modifying (4) by the introduction of a dimensionless factor c, so that it becomes

$$mg = 2\pi cr\sigma. \tag{5}$$

This coefficient was determined by them as a function of the dimensionless ratio $r/v^{\frac{1}{3}}$, v being the volume of the drop (Fig. 8).
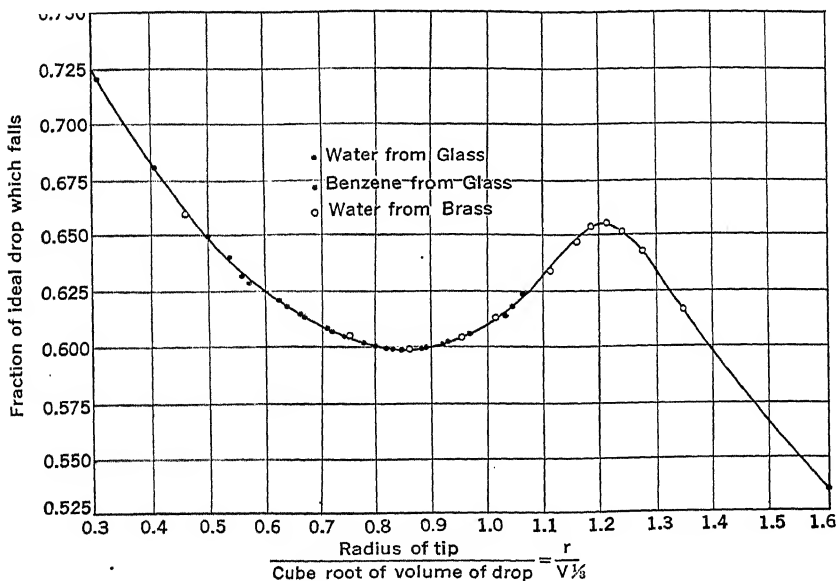
Fig. 8.   Correction Plot for the Drop-Weight Method.*

Thus, in making an experimental determination of the surface tension of a liquid by this method, the weight of a given number of drops formed *slowly* from a sharp tip of known radius is measured.   Knowing the density of the liquid, the value of $r/v^{\frac{1}{3}}$ can be calculated, and the corresponding value of c obtained from the tables prepared by Harkins.†

* HARKINS, W. D., and BROWN, F. E., *J. Am. Chem. Soc.* **41**, 499 (1919).

† It was formerly customary to calibrate tips by dropping a liquid of known surface tension and assuming drop weight proportional to the surface tension, but this is allowable only when c has the same value for both liquids.   It is clear from Fig. 8 that, using this method, errors of 10% may easily result.   An elegant method involving photographic measurement of the pendant drop itself has been developed by Andreas, Hauser, and Tucker, *J. Phys. Chem.* **42**, 1001 (1938).   It has the important advantage of making possible quantitative study of the effect of age on the interface.

The most important method of measuring surface tension (capillary rise) is complicated by the behavior of a liquid in contact with other surfaces, and description of it is therefore postponed.

Contact Angle.

If a flat glass plate AB is immersed vertically in a liquid, so that part of it projects above the surface, as shown in Fig. 9, there will be two main forces acting upon a liquid molecule, such as C, near the glass. The forces of mutual intermolecular attraction between the liquid molecules tend to drag the mole-
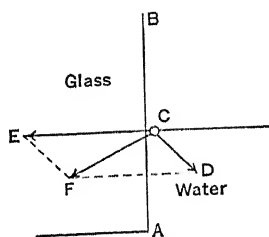
Fig. 9.
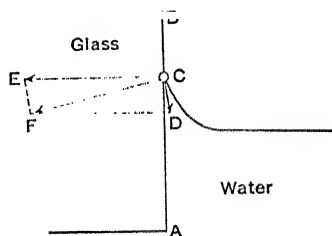
Fig. 10.
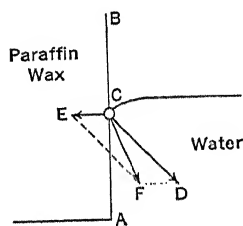
Fig. 11.

FIGS. 9–11. Solid-Liquid Interface

cule in question down into the liquid, but at the same time it is drawn towards the glass by the attraction upon it of the field of force of the latter. Imagine these forces of cohesion and adhesion respectively to have magnitudes and directions represented by the lines CD and CE. These must have a resultant CF, and the surface of the liquid will tend to set itself at right angles to it. This requires that C and the molecules adjoining it move up the plate. As this occurs, the forces acting on C due to cohesion will become more nearly vertical but far smaller in magnitude. Hence the resultant CF is more nearly perpen-

dicular to the glass. In other words, the liquid cannot remain perpendicular to the plate as shown in Fig. 9, but must form a meniscus appearing somewhat as in Fig. 10. In this latter figure, C is visualized as located at the ultimate point of contact of the water with the glass. The angle which the water surface makes with the solid at this point is known as the contact angle, and, for the particular case considered, clean water and glass, is zero.

The value of the contact angle for any particular system is dependent upon the relative attractions of the liquid molecules for each other and for the solid. If the attraction for the solid is sufficiently large, then, as has been seen, the resultant force will be nearly perpendicular to the glass wall, and the contact angle will be zero. If, however, the molecules of the liquid have relatively little affinity for the solid, e.g., paraffin wax and water, the cohesive force is so much greater than the adhesive that the resultant force CF (Fig. 11) is directed downwards into the liquid and the liquid surface assumes a downward bend. Thus, a liquid and solid may have a contact angle of any value between $0°$ and $180°$, the value depending upon the relative attractions of the solid and liquid molecules.

Accurate values of contact angles are hard to obtain experimentally because minute traces of impurities have great effect upon their value. They often differ for advancing and receding meniscuses, the cause of the irreversibility perhaps being due to changes in the surface resulting from contact with the liquid. They can be determined by immersing a plate of the solid in the clean liquid, and measuring with a protractor its angle of tilt when the liquid surface meets it horizontally without bending, since incipient curvature of the liquid surface near the plate is readily detectable.[*] Langmuir [†] has described an exceedingly simple means of finding contact angles, depending upon the determination of the angle at which a beam of light is just reflected from the surface of a drop placed upon the solid in question.

[*] The contact angle against glass may also be obtained by measuring the surface tension of the liquid, using two methods, one independent of and the other dependent upon contact angle. Thus, the maximum bubble pressure and capillary rise (pp. 40 and 46) methods might be used.

[†] LANGMUIR, I., and SCHAEFER, V. J., *J. Am. Chem. Soc.* 59, 2400 (1937).

In general, it is found that the contact angles of water and most organic liquids against glass are zero, a similar result being obtained for mercury on amalgamated copper. Mercury on steel, on the contrary, has a contact angle of about 150°; water on paraffin wax also shows a large contact angle, 107°.

The Capillary Rise Method.

The rise of a liquid up a straight glass plate has already been considered, and it is plain that the same forces will cause a liquid to rise or fall in a glass tube, the direction being dependent upon the contact angle. Around the circumference AB, of the capillary of Fig. 12, the liquid, having a contact angle less than 90°, may be pictured as clinging to the tube with a force of $2\pi\sigma$ acting along AC, or, if this be resolved in the vertical plane, $2\pi\sigma \cos \theta$. This vertical component is supporting the weight of the fluid column of height h. Thus, if ρ is the density of the fluid,

$$2\pi r\sigma \cos \theta = g\rho\pi r^2 h,$$

whence

$$\sigma = \frac{grh\rho}{2\cos\theta}. \tag{6}$$

If the contact angle is zero, $\cos\theta = 1$, reducing to

$$\sigma = \frac{grh\rho}{2}. \tag{7}$$
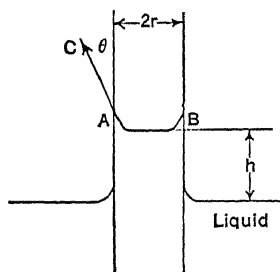
FIG. 12. Liquid Rise in a Capillary Tube.

This arrangement can be used to measure the surface tension of a liquid, although it will be seen that it is dependent upon the contact angle, which must be accurately known. For absolute determinations, the method suffers from the difficulty of obtaining tubes exactly circular in section and uniform in radius over considerable length, but is probably the most accurate available for the measurement of surface tension of pure liquids. When the tube can be calibrated against liquids of known surface tension, the method becomes highly satisfactory. Surface tensions, determined by Richards and Carver, are given in Table I, water and benzene being particularly satisfactory as standards for calibration.

TABLE I.   SURFACE TENSIONS IN AIR AT 20° C.*

Liquid	Surface Tension, in dynes/cm.
Water	72.75
Benzene	28.88
Chloroform	27.14
Ether	16.96
Carbon tetrachloride	26.77

Interfacial Tension.

When two immiscible liquids are in contact, the molecules of each liquid tend to exert an attraction upon those of the other in the surface, which somewhat lessens the inward pull of each liquid upon the molecules of its own kind present at the mutual surface.   Thus a molecule of either liquid situated in the interface will be under different forces from a similar molecule in the main bulk of the liquid.   Hence, the dividing surface will have a surface energy, as at a liquid-gas interface.   This surface energy is known as the interfacial tension, and is always less than that of the liquid with the higher surface tension.   For miscible liquids, since there is no interface, the interfacial tension must be zero.

The interfacial tension between liquids is a constant of importance.   Theoretically, all the methods used in the determination of surface tension are available for its measurement, but practically only the drop-weight method and modifications of the capillary-rise and ring methods are convenient.   Since the interfacial tension between two liquids depends upon the degree of mutual saturation, to obtain definite, reproducible results the two phases should be completely saturated with each other.

Spreading.

If a small drop of hydrocarbon oil, *e.g.*, tetradecane, is placed on the surface of water in a beaker, the droplet floats as such. If more oil is added to the drop it increases somewhat in area, decreasing in curvature, and flattens out to a floating lens. Further additions of oil increase the area with but little change

* RICHARDS, T. W., and CARVER, E. K., *J. Am. Chem. Soc.* 43, 827 (1921).

in thickness of the layer. If, however, instead of a hydrocarbon, one uses a drop of oleic acid, the behavior is extraordinarily different. The smallest drop disperses extremely rapidly over the whole surface of the water the instant it comes in contact with it. Furthermore, a very small amount of acid is sufficient to cover completely a large water area. Behavior such as that of oleic acid is called spreading, in contradistinction to the non-spreading characteristics of the hydrocarbon oil.

It is found experimentally that incompletely miscible liquids both containing groups or radicals of approximately the same polarity tend to spread on each other. If, on the other hand, the two liquids be completely dissimilar in polar character, spreading does not occur. When the two liquids are sufficiently similar in their polar character, solution will immediately follow spreading (p. 17).

The spreading or non-spreading of a liquid can also be approached from the energy viewpoint. Visualize a relatively large area of non-spreading liquid resting on the surface of a second, as shown in Fig. 13. The surfaces of the lens will be substantially flat except at the edges, but the lens is sufficiently large so that the curved edge surfaces and their effects are negligible in comparison to the whole. Call the total area of the lens A. The free energy resident in the upper surface of the lens is $\sigma_o A$; that in the lower surface $\sigma_{ow} A$. If the sum of these two free energy quantities be greater than the free energy of an equal surface of water, $\sigma_w A$, the lens can be at equilibrium only in case it rides at a level higher than that of the water surface AB, so that the energy content due to this level exactly balances the excess of the surface energies. In other words, the difference in the surface energies determines the thickness of the lens. If, now, one replace the lens by one of a second liquid for which the value of $\sigma_o A$ plus $\sigma_{ow} A$ is less, the excess height required to balance the surface energies (*i.e.*, the lens thickness) will decrease and the lens increase in area correspondingly. If the term $\sigma_o A$ plus $\sigma_{ow} A$ be decreased until it is substantially equal to or less than $\sigma_w A$, the lens will flatten out indefinitely and one will

A ——————— Oil ——————— B
Water

FIG. 13. Liquid Drop upon Water.

be dealing with a spreading liquid. The term $\sigma_w - (\sigma_o + \sigma_{ow})$, serving as it does as a criterion of spreading, has been called by Harkins the spreading coefficient S of one liquid upon another.*

Experimentally, it is found that two pure liquids sometimes spread on each other, while the same liquids when mutually saturated with respect to each other do not spread. Thus, it is important to distinguish between initial and final spreading coefficients. Many substances, *e.g.*, the alcohols and phenols on water, possess high positive initial spreading coefficients, but, when the two liquids are mutually saturated, a negative final spreading coefficient. Illustrative initial spreading coefficients are given in Table II.

TABLE II.   INITIAL SPREADING COEFFICIENTS FOR ORGANIC
LIQUIDS ON WATER AT 20° C.†

(a) *Spreading Liquids*	*Coefficient (ergs/cm.²)*
Benzene	8.86
Hexane	3.07
Nitrobenzene	3.71
iso-Butyl alcohol	48.15
Heptylic acid	37.88
Oleic acid	24.57
Chlorobenzene	2.26
(b) *Non-Spreading Liquids*	
Methylene iodide	−23.80
Carbon disulfide	− 7.26
Medicinal paraffin (average)	−13.6
Acetylene tetrabromide	−15.74

Antonow ‡ suggested that the interfacial tension between two mutually saturated liquids is equal to the difference of the sur-

* Cohesion is the force of attraction between like molecules. If, for a liquid, it be arbitrarily defined as the reversible, isothermal work necessary to separate a column of the substance one square centimeter in cross section into two parts by rupture, since there are now two square centimeters of fresh surface, the free energy change is 2σ. Harkins has defined this quantity as the work of cohesion W_c. Adhesion is the force of attraction between unlike bodies. If one visualize two rods in contact end to end, each rod one square centimeter in cross section, separation of the rods along their interface involves a free energy change of $\sigma_A + \sigma_B - \sigma_{AB}$ where σ_{AB} is the surface tension of the two materials at their interface. This term has been defined by Harkins as the work of adhesion W_a (see p. 55). From the above it is clear that the spreading coefficient is

$$S = W_a - W_c.$$

Spreading occurs spontaneously when the work of adhesion between the substances concerned is greater than the work of cohesion of the liquid forming the drop.

† *Colloid Symposium Monograph* 6, 17, Chem. Cat. Co., Reinhold Publishing Corp. (1928).

‡ ANTONOW, G. N., *J. Chim. Phys.* 5, 372 (1907); *Koll. Z.* 59, 7 (1932); 64, 336 (1933).

face tensions of the independent liquids. Although this holds in a number of cases (Table III), it does not hold for non-spreading liquids nor for those having a high initial spreading coefficient upon the second liquid.*

TABLE III.   INTERFACIAL TENSIONS OF LIQUID-WATER SYSTEMS AT 25° C.†

Liquid	S. T. Aq. Phase	S. T. Org. Phase	I. T.	Deviation
Benzene	61.8	28.36	33.64	- 0.2
Nitrobenzene	68.00	43.05	24.81	0.14
Methylene iodide	71.86	52.30	40.52	-20.96
n-Butyl alcohol	25.23	24.63	4.04	- 3.44
Diethyl ketone	35.65	26.10	9.20	0.35
m-Cresol	37.27	37.78	3.60	- 4.11

In the case of the spreading of liquids upon solids, although a similar relation to that for two liquids presumably applies, it has little practical utility because the surface tension of the solid and the interfacial tension between the liquid and the solid are not subject to dependable measurement. If, on the other hand, a given liquid has a finite contact angle with a specific substance, it is clear that a drop of the liquid placed on a horizontal surface of the solid will not spread indefinitely but will form a lens of finite thickness. If the contact angle be sufficiently high, the lens becomes a distorted spheroid, with a re-entrant angle at contact with the plane, such as the sessile drop formed by mercury on glass. However, liquids of zero contact angle obviously tend to spread indefinitely on a horizontal surface. Indeed, in many cases such liquids have sufficient affinity for the solid to rise almost indefinitely up vertical surfaces, and are said to wet the solid (see p. 76).

Surface Tension and Temperature.

As previously pointed out, the kinetic agitation of the molecules increases with rise in temperature, while the attractive forces between the molecules, though changing little specifically,

* It is seen that for the rule to hold, the final spreading coefficient S' (i.e., the spreading coefficient for the mutually saturated liquids) must be zero; for, since $\sigma_w - \sigma_o = \sigma_{ow}$,

$$S' = \sigma_w - (\sigma_o + \sigma_{ow}) = 0.$$

† CARTER, E. G., and JONES, D. C., *Trans. Faraday Soc.* 30, 1027 (1934).

are reduced in effect on account of the decrease in density with rise in temperature. Surface energy should therefore decrease with rising temperature, and for most compounds does in fact fall almost linearly. In view of equation (1), this would require that the total surface energy be independent of temperature. However, while surface energy becomes zero at the critical temperature, it ceases to be linear in its neighborhood. Ramsay and Shields suggested the relation

$$\sigma = c\,(T_c - T - K) \qquad (8)$$

where T_c is the critical temperature and c and K constants. This breaks down near the critical temperature, a better relation being that of Ferguson

$$\sigma = \sigma_0 \left(1 - \frac{T}{T_c} \right)^n \qquad (9)$$

where n averages about 1.21, usually lying between 1.192 and 1.248.

BIBLIOGRAPHY

ADAM, N. K., *The Physics and Chemistry of Surfaces*, Ed. 2, Clarendon Press, Oxford (1938).

ANDREAS, J. M., HAUSER, E. A., and TUCKER, W. B., *J. Phys. Chem.* **42**, 1001 (1938).

BARTELL, F. E., " Wetting Agents," *Chem. and Industry* **60**, 475 (1941).

BURDON, R. S., *Surface Tension and the Spreading of Liquids*, Macmillan (1941).

CHAMPION, F. C., and DAVY, N., *Properties of Matter*, Prentice-Hall (1937).

CLAYTON, W., *Theory of Emulsions*, Ed. 3, Blakiston, Philadelphia (1935).

DORSEY, N. E., "Measurement of Surface Tension," *U. S. Bureau of Standards, Scientific Papers*, No. 540, Bull. **21**, 563 (1926).

HARKINS, W. D., "Surface Energy and Surface Tension," in Alexander, J., *Colloid Chemistry*, Vol. I, 192, Chem. Cat. Co. (1926).

Chapter IV

Surface Tension and Orientation

Attention has already been called to the fact that, whereas the free surface energies of liquids change relatively rapidly with the temperature, their total surface energies are almost constant and independent of temperature.* Granting that one wishes to compare the surface energies of different liquids, there is no obvious reason for comparing them at equal temperatures and the choice of any one temperature of comparison is open to dispute. An obvious way out of the dilemma is to compare total surface energies. When this is done certain extraordinary relationships come to light. Thus, the total surface energies of saturated straight-chain paraffin hydrocarbons are relatively independent of molecular weight, having values of approximately 50 ergs per sq. cm. Not only this, but the surface energies of the corresponding primary alcohols have approximately the same value. Comparison of the surface energies of the corresponding fatty acids gives values slightly higher, approximately 55 ergs per sq. cm., but even here the small difference in this characteristic liquid property for these distinctly different compounds is striking. The same is true for the primary amines with values of approximately 52 ergs per sq. cm.

There must be an underlying explanation of the limited variation of the surface energies of such widely different compounds. The figures force one to conclude that the surfaces of the liquids themselves must possess some underlying identity in physical structure. Keeping in mind the fact that all these compounds, other than the hydrocarbons, are characterized by the presence in the molecule of a highly polar radical, whereas the hydrocarbons themselves are decidedly non-polar, it appears probable that the surface even of the polar compounds has the

* See equation (1), p. 38.

structure of the hydrocarbons, essentially unaffected by the polar group.   Otherwise it is difficult to explain the identity of the surface behavior of the two classes.   This is the more plausible

TABLE I.   TOTAL SURFACE ENERGIES OF SOME ALIPHATIC
COMPOUNDS *

Name	Formula	Total Surface Energy, ergs/sq. cm.
Hexane	C_6H_{14}	49.5
Octane	C_8H_{18}	48.4
Ethyl alcohol	C_2H_5OH	46.4
n-Octyl alcohol	$C_8H_{17}OH$	50.8
Acetic acid	CH_3COOH	55.9
Butyric acid	C_3H_7COOH	54.5
iso-Caproic acid	$C_5H_{11}COOH$	50.2
Ethylamine	$C_2H_5NH_2$	48.1
n-Propylamine	$C_3H_7NH_2$	50.4
Water	H_2O	118.1
Chloroform	$CHCl_3$	59.7
Carbon tetrachloride	CCl_4	62.2

in view of the fact that the polar groups, because of their strong fields of force, would tend to bury themselves in the liquid, whereas the non-polar portions of the molecule should have this tendency in far less degree.   In other words, the hypothesis suggests itself that in the polar compounds the molecules in the surface layer may be oriented, the polar groups inwardly directed, and the surface thus consist of the hydrocarbon chains (Fig. 1).†
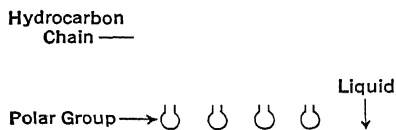
Fig. 1.   Orientation in the Interface.

The surface energy is essentially that of orienting the hydrocarbon chains more or less side by side in the surface.   This substantial parallelism of the long hydrocarbon chains would resemble a closely packed pile structure.

* HARKINS, W. D., DAVIES, E. C. H., and CLARK, G. L., *J. Am. Chem. Soc.* **39**, 541 (1917).

† There cannot be complete orientation, since the surface molecules are in continuous thermal agitation, their estimated average life in the surface being of the order of 10^{-7} sec. This is confirmed by the fact that substituents, which must influence both orientation and agitation, do have some effect upon the total surface energy, which increases slightly as H is replaced by $- OH$, $- NH_2$, $- CN$, $- COOH$, $- NO_2$, etc.

Orientation at Liquid-Liquid Surfaces.

Further light is thrown on the structure of liquid surfaces by study of liquid-liquid surface energies. The free surface energy of a liquid against its vapor (assumed dilute) is the work required to bring the molecules from the interior to a zone on one side of which molecular attractions are negligible. On the other hand, a molecule coming into the interface of two immiscible liquids is entering a zone where the fields of force are still considerable; hence the energy required must be less, *i.e.*, the interfacial energy is less than the sum of the surface energies of the two components, and may be estimated tentatively as one half this value. For the paraffins and water this is approximately true (Table II), but wide deviations are obtained for polar compounds such as heptylic acid, the observed value being much lower than that thus estimated. This demonstrates that much less energy is required for the formation of an acid-water than for a paraffin-water surface.

The following reasoning, due to Harkins, leads to a similar result. Imagine a bar of a given liquid of unit cross-sectional area to be separated into two portions, without change in cross-sectional area. To form the new surfaces work will have to be expended. Harkins calls this the work of cohesion; it is obviously equal to 2σ, if σ is the free surface energy or surface tension of the liquid. Similarly, the work of adhesion between two immiscible liquids can be imagined as the work necessary to separate two bars, of unit cross-sectional area, of the liquids joined together at their interface. In this way a surface of each liquid against air will be formed, in place of the one interface between the two liquids. The work expended in separating the two liquids is the work of adhesion, given by

$$W_a = \sigma_A + \sigma_B - \sigma_{AB}. \tag{1}$$

Measurement of the surface tensions and interfacial tensions (by the methods given in Chapter III), therefore, offers a measure of the attractive forces between the molecules of different liquids, and a study of such measurements discloses significant relations. In Table II are given values for the surface tensions and interfacial tensions against water of a number of organic liquids,

from which have been computed the values for $W_a - W_c$, which, it will be remembered (Chapter III, p. 49), is the spreading coefficient.   It is seen that the adhesion between n-octane and

TABLE II.   WORKS OF ADHESION AND COHESION TO WATER
OF VARIOUS ORGANIC LIQUIDS AT 20° C.
$$\sigma_W = 72.8 \text{ ERGS/CM.}^2 *$$

Liquid	σ_L	σ_{LW}	$\dfrac{\sigma_W + \sigma_L}{2}$	W_c	W_a	$W_a - W_c$
Hexane	18.4	51.2	45.6	36.8	40.0	3.2
Octane	21.8	50.8	47.3	43.6	43.8	0.2
Di-iso-amyl	22.2	46.8	47.5	44.4	48.2	3.8
Octyl alcohol	27.5	8.5	50.2	55.0	91.8	36.8
iso-Valeric acid	25.3	2.7	49.1	50.6	94.6	44.0
Heptylic acid	28.3	6.6	50.6	56.6	94.8	38.2
Nitromethane	36.8	9.7	54.8	73.6	99.9	26.3
Benzene	28.9	35.0	50.9	57.8	66.6	8.8
Toluene	29.9	36.1	51.4	59.8	66.6	6.8
Aniline	42.6	5.8	57.7	85.2	109.6	24.4

water is about the same as the cohesion for octane itself; for octyl alcohol, on the other hand, the adhesion against water is much greater than the cohesion of the alcohol.   In general, the adhesion for hydrocarbons against water seems to be about 50 ergs per sq. cm., while for compounds containing a polar group, $e.g.$, hydroxyl or carboxyl, it is much higher, of the order of 90 ergs per sq. cm.   In both series the adhesion is nearly independent of chain length.   Certain groups, such as the amino and carboxyl groups, because of their marked affinity for water, appear to confer on molecules water solubility or a tendency thereto. Orientation provides a clear explanation of the difference between the energies of cohesion and adhesion. If, at the interface between alcohol and water, the molecules of octyl
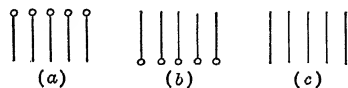
FIG. 2.   Orientation of Polar and Non-Polar Molecules in the Surface.

alcohol were arranged as in Fig. 2(a), since the hydrocarbon chains are presented to the water as in an octane-water interface, the work of adhesion would be expected to be substantially

* Data of HARKINS, W. D., CLARK, G. L., and ROBERTS, L. E., $J.$ $Am.$ $Chem.$ $Soc.$ 42, 700 (1920).

the same as for the hydrocarbon [Fig. 2(c)]. Actually, it is much larger, pointing to some such arrangement as Fig. 2(b), in which the molecules are oriented with their hydroxyl groups attracted towards and immersed in the water, as is to be anticipated in view of the polar character of the groups. This indicates that separation of an organic liquid containing a polar group from water requires more work, because polar groups have to be separated from each other. $W_a - W_c$ should, therefore, be a measure of the polarity of the organic liquid. It does, in fact, parallel the other available measures of polarity (p. 16).

Surface Films.

If, instead of considering the behavior of two substances when brought together in bulk, attention is focused on the phenomena occurring when a small amount of a substance is dropped onto a large surface of a second material, still further confirmation of the orientation theory is obtained and the picture of molecular forces suggested in the previous pages receives striking verification. Suppose a small amount of a hydrocarbon or some other substance with no affinity for water is placed on a water surface; there is, as already discussed on p. 47, no spreading. The drop floats on the water as a circular lens, the precise shape of which is dependent upon the surface tensions and the densities of the two liquids, since gravity is also acting upon the drop. Equilibrium is invariably reached while the lens is still thousands of molecules thick. On the other hand, if a minute amount of some substance containing a group with sufficient affinity for water, e.g., pentadecylic acid, $C_{14}H_{29}COOH$, be placed on a water surface at room temperature, the substance will spread but will not dissolve, since the presence of one carboxyl group is insufficient to render the whole molecule water soluble. Because powders, such as talc, cling to films of the acid, it can easily be shown, by dusting talc on the surface, that the acid molecules spread over the water without apparent limit, in marked contrast to the behavior of a hydrocarbon such as tetradecane, $C_{14}H_{30}$. Clearly, the acid is being dispersed over the water surface as an excessively thin film by some sort of expansive force which it should be possible to measure.

The variation in the pressure exerted by the molecules in a surface film with its changes in area can be investigated by means of the surface balance, first devised by Langmuir* and elaborated by Adam. It consists essentially of a shallow brass trough (Fig. 3) filled to the brim with water, over which a barrier *AB* can be moved. In this way the area *ABEF* of the film can be adjusted at will. Leakage past the barrier is prevented by coating both it and the sides of the trough with hard paraffin wax, which does not itself contaminate the water surface. *CD* is a very light, movable, copper barrier extending to within approximately a millimeter of each side, kept in position by
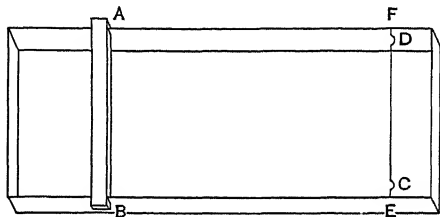
FIG. 3. Surface Balance.

a force transmitted from a fine torsion wire (mechanism not shown). The force applied to the barrier is found by previous calibration of the torsion wire. Movement of the barrier is indicated by a mirror and scale or other device, so that it can be kept in any desired position with considerable precision. Two thin, flexible, platinum ribbons *CE* and *DF* prevent leakage of the film past the ends of *CD*. The surface of the water is first tested for any accidental contamination by bringing the barrier *AB* up towards the float; if no impurities are present the pressure registered on the float will be almost negligible. The substance under investigation is then placed on the surface in *ABEF*, usually by dissolving it in a volatile solvent, immiscible with water (*e.g.*, light petroleum), and dropping a known volume of the solution onto the surface, the evaporating solvent leaving the substance to form a film. The area *ABEF* can then be changed as desired by moving the barrier *AB*, and the corresponding force on *CD* determined from the reading of the torsion wire. In this way the force or pressure exerted upon the barrier by the molecules in the film can be measured to 0.01 dyne per cm.

When the magnitude of the expansive force of the film exerted against the barrier *FE* of Fig. 3 is measured, it is found

* LANGMUIR, I., *J. Am. Chem. Soc.* 39, 1848 (1917).

that, for many substances, it drops off progressively as the area of the film is increased. Plotting the force against the area, at relatively large areas, gives a curve of the type shown in Fig. 4, the force becoming progressively smaller as the area increases.
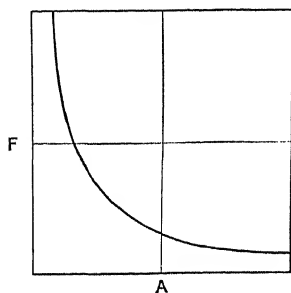
Fig. 4. Force-Area Curve for a Gaseous Film.

In other words, there exists on the surface a structure capable of continuous expansion to an unlimited extent, provided the restraining force is sufficiently reduced. The only known form of matter of which this behavior is characteristic is a gas; that is, the film, *e.g.*, of acid, on the surface of the water has the expansive properties of a gas. The force-area curve of Fig. 4 is analogous to the pressure-volume curve for isothermal expansion of a gas or vapor.

Study of the situation makes it clear, however, that the acid film on the water surface is no ordinary gas. Whereas a gas can expand freely in all directions, the acid, because of its low volatility, cannot expand upward into the vapor space, nor, because of its insolubility in water, downward into the liquid. It is capable only of lateral expansion in the surface.

Two-dimensional expansion of a gas may be visualized by imagining the gas to be confined between two parallel planes, sufficiently close together so that, on the average, only a single molecule of the gas exists at any particular point between the two surfaces. Considering 1 gram mol of the gas,

$$PV = RT.$$

Apply this equation to a single molecule,* $PV = RT/N$. Visualizing the volume V occupied by this individual molecule as a cube of side L, $RT/N = PL^3$. L, on the assumption employed, constitutes the distance between the two restraining planes.† Consider a surface between the planes and perpendicular to them, whose length is unity. Since its width is L, its area is also L. The force exerted upon this surface by the gas is the

---

* Obviously, this implies a statistical average pressure effect.
† It is also the distance below which the restraining action of the planes would prevent the gas from continuing to follow the gas laws.

pressure times the area, or PL. Calling this force per unit
length F, $RT/N = FL^2$. However, per individual molecule, L^2
is the area A occupied by it, measured on the confining planes.
Hence,

$$\frac{RT}{N} = FA. \tag{2}$$

Expressing volumes in c.cm. and pressures in dynes per sq. cm.,
R becomes $8.315 \cdot 10^7$ ergs per ° K. per mol, whence

$$FA = 1.37 \cdot 10^{-16} T, \tag{3}$$

or, at $t = 20°$ C., $FA = 4.01 \cdot 10^{-14}$. The constant is rendered
convenient in magnitude by expressing the area per molecule in
square Ångström units,* whence

$$FA = 400. \tag{4}$$

Granting that the molecules in the film of acid on the surface
of the water are free to move on the surface, possess negligible
mutual attractions except perhaps in the immediate neighbor-
hood of each other, but are so constrained that they cannot
possibly leave the surface, these molecules should behave as a
gas conforming substantially to the equation just derived. If
so, the curve of F vs. A (Fig. 4) should be a rectangular hyper-
bola whose constant, expressing F in dynes per cm. and A in
square Ångström units, should be 400. The experimental data
in fact show this to be the case.† In other words, these surface
films conform at low pressures both qualitatively and quantita-
tively to the gas laws, on the assumption that the molecules in
the film are free to move in any direction in the surface but
incapable of effective movement at right angles to it. Because
of these characteristics such films have been described as two-
dimensional gases.

The analogy between the behavior of the surface film and a
gas extends much further. Thus, at elevated pressures the iso-

---

* One Ångström unit $= 10^{-8}$ cm. Since the amount of substance placed on the
surface of known area, its molecular weight, and the Avogadro number are known, the
area occupied per molecule (A) can be calculated readily.

† Experimentally, the theoretical value of 400 is only approached at very low pres-
sures. The value to which the FA product tends at zero pressure is difficult to determine
exactly since the pressures measured are very small, fractions of a dyne in most cases.
The theoretical value of 400 is checked, however, to within at most 10%.

thermal PV product for a gas is no longer constant, but varies with the volume. Similarly, for these gaseous surface films the FA product is not constant but varies with the area, as shown for gaseous films of dibasic esters ($EtOOC \cdot (CH_2)_n \cdot COOEt$) in

Fig. 5. Furthermore, the shape of this FA vs. A curve is similar to that of the PV vs. V curve for a gas (see Fig. 5, Chapter I). Both show that cohesion between the molecules is always present and is responsible for the initial dip in the FA vs. A curve, just as intermolecular attraction is responsible for the similar dip in the PV curve of a gas. As the pressure increases, however, the area occupied by the molecules themselves becomes important and the FA product becomes large, exactly as the PV product does in the case of a gas.

Fig. 5. FA–F Curve for Dibasic Esters.*

Gaseous films are not common since they exist only at low pressures and high surface areas. Indeed they can be formed at measurable pressures only if lateral cohesion between the chains is sufficiently low. Because of the precision required in making measurements in the region of low pressures their existence was long overlooked, just as the existence of gases in the three-dimensional world was overlooked until about the sixteenth century. At higher pressures, where the experimental technique is less difficult, by analogy liquid and solid films might be expected to exist and these, actually, were the first to be discovered.

Liquid Films.

If the surface area of a suitable gaseous film is diminished by bringing the sliding barrier closer to the pressure barrier, the force exerted on the latter rapidly becomes less than that required by (2), remains constant over a considerable range (Fig. 6), then increases very rapidly, until finally the film begins to crumple and visible strain lines appear on the surface. In this final stage the monomolecular layer no doubt begins to give way and the molecules are squeezed out.

* Adam, N. K., and Jessop, G., *Proc. Roy. Soc.*, A, 112, 376, Cambridge University Press, Macmillan (1926).

As the area is decreased from A to B, apparently the molecules are pushed closer together and the forces of attraction between them become operative, causing "liquefaction" to commence at the point B. The co-
hesive force between the mole-
cules is sufficiently strong from
B onwards to cause formation of
clumps of molecules or islands
of film, in which a group of
molecules stick together while
surrounding molecules still move
around freely. This inhomo-
geneity of the film can be demon-

Fig. 6. Force-Area Curve during Liquefaction.

strated by surface potential measurements * or other methods. The portion BC of the curve therefore represents the "liquefaction" of the gaseous film. At C this has become complete and a coherent film is obtained. The constant surface pressure exerted by the film on the barrier from B to C is analogous to the vapor pressure of a liquid and may be called the film vapor pressure.† From the rapid rise in the curve from C onwards, it appears that the film is being compressed until finally it crumples and gives way under the stress. The similarity to the liquefaction of a gas becomes even more striking when the family

---

* A metal electrode above but near the surface of a liquid is found to possess a potential relative to the liquid, measurable, *e.g.*, against a standard half cell in the liquid, provided the air gap is rendered conducting by ionization. The magnitude of the potential depends on the state of the surface. Thus, the method makes it possible to identify lack of homogeneity in surface films, or transition from one type to another.

† It is of interest to calculate this in equivalent conventional units, assuming for the moment that the film has a finite thickness.

$$FA = \frac{RT}{N} \quad \text{or} \quad FL^2 = \frac{RT}{N},$$

whence

$$P = \frac{RT}{NL^3} \quad \text{or} \quad \frac{RT}{NA^{\frac{3}{2}}},$$

where P is the equivalent three-dimensional pressure. Hence

$$P = \sqrt{\frac{NF^3}{RT}}.$$

At 14.5° C. the "vapor pressure" of pentadecylic acid is 0.11 dynes per cm. (Fig. 7), whence, from the formula just given, $P = 135$ mm. mercury. Pentadecylic acid is a solid at 14.5° C. and its vapor pressure is negligible, indicating that the attraction of its molecules for each other in the solid state has been largely overcome in the monomolecular film on the surface of the water.

of curves (Fig. 7) for the various aliphatic carboxylic acids from C_{12} to C_{16} is considered. These decrease in volatility with rising molecular weight. From the curves lauric acid is above the "critical point," sufficient lateral cohesion between the molecules to form a coherent film being absent. In palmitic acid,
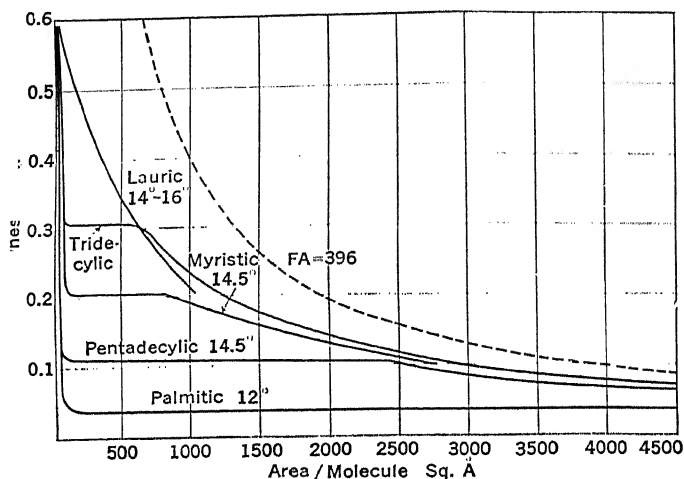
FIG. 7.   Force-Area Curves for the Aliphatic Acids.*

on the contrary, the cohesion between the molecules is so great that the acid never vaporizes completely at pressures high enough to be within the experimental accuracy of the apparatus.

Condensed Films.

Films, in which individual molecules are not free to move of themselves, but cohere strongly and form islands or compact masses on the surface, are known as condensed films. They may be liquid or solid, the two states being distinguishable by scattering talc over the surface of the film and blowing gently. If the film is liquid, the powder moves freely, but if solid, it is immobilized. Generally, compounds with long carbon chains, particularly if the chain contains over 20 carbon atoms, give solid films, which may have considerable rigidity. Since the condensed films exert relatively high pressures at small areas,

they were the first to be investigated. Langmuir,* in a classical research, obtained very significant results from their study.

The relation between area and pressure has been studied with the surface balance for a large number of compounds in the region of 0 to 60 dynes per cm. pressure and 0 to 40 sq. A area, roughly corresponding to the portion CD of the curve in Fig. 6. Figure 8 shows the F–A curves on water for three typical homologous series; ABC for long-chain saturated fatty acids, ABD for the corresponding esters, and EF for para-substituted phenol derivatives in which the group $-C_6H_4OH$ replaces the carboxyl group of the fatty acids. The corresponding curves for many analogous aliphatic derivatives giving similar films can be divided, like ABD, into two sections, a portion DB at relatively low pressures and a portion AB at higher pressures. For others, as with the fatty acids on distilled water, DB is merely a continuation of AB. A number of important generalizations can be made from the results. First, for a given homologous series of the types shown in Fig. 8, the F–A relation is independent of the number of carbon atoms in the aliphatic chain, provided this be long enough to confer substantial insolubility in water. Second, the upper part of the curves (the portion AB of Fig. 8) is the same for many straight-chain, aliphatic α-derivatives having a monovalent head group, irrespective of the structure of that group, e.g., the amines, methyl ketones, and esters of carboxylic acids. Third, different head groups frequently give characteristic lower parts (the range BD) of the curves, but the break-away of BD from AB, when present, is always in the direction of greater rather than lower surface areas.

While the curves AB and EF of Fig. 8 are nearly vertical,
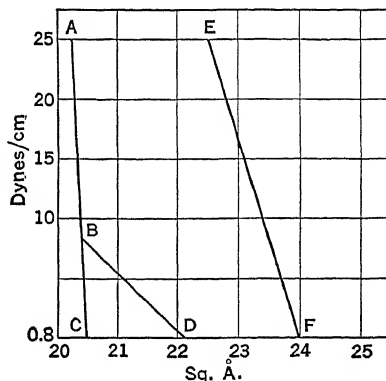
FIG. 8. Force-Area Curves for Condensed Films.†

* LANGMUIR, I., *J. Am. Chem. Soc.* **39**, 1848 (1917).

† Data from ADAM, N. K., *The Physics and Chemistry of Surfaces*, The Clarendon Press, Oxford (1938).

they do in fact slope downward to the right. The intercepts of these curves extrapolated to zero pressure presumably give the cross-sectional area of the molecules in the film.* This area is 20.5 sq. Å per molecule for an extraordinarily large number of α-straight-chain derivatives, as stated above. This is true in compounds actually measured, despite variation in the hydrocarbon chain length of from 13 to 34 carbon atoms. It seems certain, therefore, that the molecules in condensed films are oriented, the active polar groups being directed into the water surface while the chain axes are substantially perpendicular to it (as in Fig. 1). Their relatively close packing is assisted by the cohesion between the chains due to their mutual molecular attractions. Because the corresponding hydrocarbons never form surface films, one must infer that existence of such films depends upon the presence of the polar groups on the ends of the chains, these groups being effective only because of their attraction for the water, causing them to be buried in the water surface. Since, however, for so many polar groups the surface area, 20.5 sq. Å, is completely unaffected by the nature of the group, this area must be a characteristic of the hydrocarbon chain itself and is presumably its effective cross section perpendicular to its length. Assuming that the density of the material in the film is identical with that of the substance in bulk, the length of the hydrocarbon chain can also be computed. Thus, the molecular volume of palmitic acid is 300 c.cm.; dividing this by the Avogadro number, the volume per molecule is 495 cu. Å. Since the cross-sectional area is 20.5 sq. Å, the length is 24.2 Å. In other words, the molecule is several times longer than it is thick, since, assuming surface arrangement of the molecules at the vertices of squares, its thickness would be $\sqrt{20.5}$, or 4.5 Å.

For more complex molecules the uncompressed area is greater. Thus, the triglycerides exhibit an area of 63.0 sq. Å, practically three times that of a single hydrocarbon chain. Even change in the head group alone is sufficient to modify the area. This is obvious from curve EF of Fig. 8, the high intercept of which is

* When closely packed but uncompressed by other than their own mutual attractive forces.

presumably the cross-sectional area of the benzene molecule, 24 sq. Å. It is noteworthy that no compounds studied have given molecular areas less than the value of 20.5 sq. Å, corresponding to straight-chain compounds. From our knowledge of molecular structure this result is to be expected.

Compounds showing discontinuity in the lower portion of the F–A curve, illustrated typically by ABD of Fig. 8, evidently exhibit this anomalous behavior because of the influence of the polar group on the end of the chain, since the location of the portion BD varies from group to group, although constant for a given group. The area D evidently represents the uncompressed close packing of the active head groups. However, the facts that these areas are often larger than might be anticipated from knowledge of molecular dimensions and that the effect frequently disappears completely under relatively low pressures (i.e., above the point B of Fig. 8), indicate that the head groups tend to occupy an abnormally high area unless restrained by external forces. This may conceivably be due to solvation (pp. 170–171).

X-ray studies on the crystals of the fatty acids indicate a cross section of the hydrocarbon chain of 18.4 sq. Å, corresponding to decidedly closer packing of the chains in the crystal than in the condensed film on the water surface. A similar result is obtained in benzene compounds, which in the crystalline condition show a ring area of 21.5 sq. Å, instead of the 24 sq. Å in the film. These figures offer an explanation of the high equivalent vapor pressure of the material in the film, estimated on page 61, in that the greater intermolecular distances in the film so weaken the mutual molecular attractions that disgregating forces have far greater effect in the film than in the crystal.

As the chain lengthens, the mutual molecular attraction of the hydrocarbon tails increases. However, the polar groups in the water have a positive tendency to separate, as demonstrated by the solubility of their corresponding short-chain derivatives. The longer the chain, the less the net tendency to expand. Stearic acid, for example, forms no gaseous film under measurable conditions (Fig. 7), whereas "vaporization" is detectable in pentadecylic, and occurs at progressively higher pressure for

acids with shorter chains, till they finally become soluble. The nature of the mutual molecular attractions is further clarified by study of the ethyl esters of the dibasic straight-chain acids, $EtOOC \cdot (CH_2)_n \cdot COOEt$. The ester groupings at the two ends of the molecule would be expected to tend to bury themselves in the water. This would necessitate the assumption that the hydrocarbon chain between these two groups lies relatively flat on the surface or else perhaps arches above it. In any event, this orientation is one in which it is difficult to visualize mutual sidewise attractions of any great magnitude between the ester molecules in the surface layer. In fact, these esters give characteristic gaseous films which remain uncondensed at relatively high pressures. However, under sufficient pressure, provided there are enough carbon atoms in the chain (17 or over), they form condensed films, which surprisingly have the normal area, 20.5 sq. Å per molecule.

Expanded Films.

Surface films show a curious state with no analogue in the three-dimensional world. The film area of many materials under a given surface pressure remains approximately constant as the temperature is raised, then, over a range of about 10° C., the area rapidly increases, finally again becoming nearly constant. Langmuir * has suggested that in this "expanded" state, the polar groups behave as if in a gaseous film, free movement being prevented, however, by the mutual attractions of the hydrocarbon chains. The heads and tails therefore behave as separate entities, the film being duplex in character.

Films on Solids.

Langmuir and Blodgett † have devised an ingenious and useful technique for transferring monomolecular films from water to a glass surface. A glass slide is dipped into water ‡ covered with a film, *e.g.*, of stearic acid, and slowly withdrawn, the surface pressure being kept constant. With each insertion and withdrawal an additional molecular layer is taken up. Films containing over 200 molecular layers have been built up and exhibit interesting properties.

* LANGMUIR, I., *J. Chem. Phys.* **1**, 756 (1933).
† BLODGETT, K. B., *J. Am. Chem. Soc.* **57**, 1007 (1935).
‡ The water should be slightly alkaline and contain calcium salts.

Soluble Films.

When the hydrocarbon chain is short and the active head has a strong affinity for water, the substances may be completely soluble, *e.g.*, propionic acid, C_2H_5COOH. However, even in this case a surface film is not absent. In a mixture of two molecular species, which have intermolecular fields of force of differing intensities, the molecules having the lower fields of force should be forced into the surface, since in this way the surface energy is reduced to a minimum. These surface forces, tending to concentrate the material in the surface at a higher concentration than that existing in the bulk of the liquid, will always be opposed by the diffusional forces, tending to make the concentration uniform throughout the solution. At equilibrium there will be a balance between the surface and diffusional forces. Thus, substances raising the surface tension of a solvent (*i.e.*, substances possessing higher fields of force than those of the surrounding molecules) will tend to be removed from the surface and become more concentrated in the solvent. The reverse will be true for substances lowering the surface tension, which will concentrate in the interface, since in this way the energy of the system is reduced.* The soluble aliphatic acids, for example, lower the surface tension of water more strongly as the acids increase in molecular weight (*i.e.*, as the molecule as a whole becomes less polar and its surface fields weaker). Thus, the acid concentrates in the surface, forming a "soluble film," in a real sense analogous to the surface films here discussed. The quantitative relation, originally derived by Gibbs in modified form, is

$$\Gamma = - \frac{c\,d\sigma}{RT\,dc}. \tag{5}$$

In this the rate of change of the surface tension σ with the bulk concentration of the solution c is related to the molal surface excess Γ of the solute per unit area of the interface. The equation states that surface concentration is proportional to the differential of the surface tension with respect to the bulk concentration. If the surface tension of the solution falls with

* This conclusion is not restricted to the two-phase system, liquid-gas. It can hold also for the systems liquid-liquid, liquid-solid, and gas-solid.

increasing concentration, then the differential becomes negative and the surface concentration of the solute will be larger than that in the bulk of the solution.   The converse holds when the differential has a positive sign.   Change in concentration of material in the surface is known as adsorption and almost invariably occurs at the interface between two phases.

Experimental verification and quantitative use of Gibbs' equation are difficult,* owing to uncertainties in determination of the surface concentration Γ.   There is no doubt of its qualitative validity, and it forms a valuable guide to the character of the adsorption in all cases to which it can be applied.

BIBLIOGRAPHY

ADAM, N. K., *The Physics and Chemistry of Surfaces*, Ed. 2, Clarendon Press, Oxford (1938).

ALEXANDER, A. E., *Trans. Far. Soc.*, **38**, 54, 1942.

* McBAIN, J. W., and HUMPHREYS, C. W., *J. Phys. Chem.* **36**, 300 (1932); McBAIN, J. W., and SWAIN, R. C., *Proc. Roy. Soc.* **A154**, 608 (1936).

Chapter V

Adsorption

At the end of the last chapter it was pointed out that a solute, if present in the liquid, exists at the vapor-liquid interface in a different concentration from that in the main bulk of the liquid. An analogous phenomenon is found at the surface of contact of gases and solids. If a gas be admitted to a highly evacuated vessel containing a non-volatile solid, provided no chemical reactions occur and the temperature be kept constant, the pressure measured is less than that calculated from the equation of state and the known dimensions of the vessel, indicating that some gas molecules must have been held by the solid surface. Visualize such a surface BC (Fig. 1); in accordance with the ideas developed in Chapter I, pp. 3–4, a molecule A of the solid will have its intermolecular forces of attraction engaged on the solid side of its sphere of action, but on the gaseous side

FIG. 1. The Gas-Solid Interface.

they will be unsatisfied. Without doubt solids, as well as liquids, possess surface energy, but whereas in a liquid surface the molecules are in active movement, in a solid surface they are anchored and each can only oscillate about a fixed point. The gas molecules, in contrast to those of the solid, are free to move in all directions at speeds governed by the Maxwell distribution law (p. 1), from time to time bombarding or colliding with the solid surface BC. As long as the incident velocity of the gas molecules is sufficient, they will hit the surface and rebound with a kinetic energy so great that the unsatisfied fields of force of the molecules of the solid surface will be insufficient to hold them. On the other hand, a more slowly moving molecule will be retained by the solid surface. Gas molecules held by the

solid may acquire sufficient energy from impacts and the oscillations or vibrations of the solid molecules to break away again or "evaporate." In this way an equilibrium is set up, the surface molecules of the solid, because of their incompletely neutralized molecular fields,* tending to hold the gas molecules on the surface. Obviously, gas molecules which have strong fields of force (*i.e.*, the more condensable gases) should be the more readily retained by the surface. This characteristic behavior, the concentration of a substance at a surface or an interface, is known as adsorption. The gas taken up by the solid surface is known as the adsorbate and the solid itself as the adsorbent.

Equilibrium is reached rapidly in many cases. Thus, active charcoal at liquid air temperature will reduce the pressure of a gas from atmospheric to a few millimeters of mercury in a few seconds. Sometimes a slow, persistent taking up of gas occurs over long periods, no doubt due to chemical reaction, diffusion into the interior with the formation of solid solutions, or other causes.† McBain suggested the name "sorption" to cover all the phenomena, often difficult to distinguish experimentally, occurring when a solid is placed in another phase in the absence of obvious chemical action ‡ or solution of the solid but with resulting change in concentration. The term "adsorption" is reserved for such changes in concentration at the interface or phase boundary and "absorption" for changes occurring in the interior of the solid. In practice, adsorption must be correlated by the use of empirical or semi-empirical equations based on experimental measurements.

The Measurement of Adsorption.

A general method for determining adsorption of gases by solids is the measurement of the change of pressure on sorption.

* Since the surface of a solid is never exactly plane, there will be some solid molecules which have larger unsatisfied fields of force than others. These will behave as "active centers." Edges and corners will also be particularly active and amorphous solids (in which the molecules are arranged at random with consequently greater unsaturated fields of force) will be more reactive than crystalline solids.

† Thus McBain, in experiments lasting over a period of eleven years, found that the amount of iodine taken up by charcoal from solution under what were apparently equilibrium conditions doubled in that time without chemical reaction.

‡ For cases involving chemical change, see pp. 76–79.

A bulb containing a known weight of the adsorbent is heated and evacuated (de-gassed).   A known quantity of the gas under investigation is introduced, brought to equilibrium with the adsorbent at a known temperature, and the pressure measured. Knowing the volume of the bulb less that of the adsorbent (dead space), one computes the quantity of gas in the bulb as such, the amount adsorbed being obtained by difference.   The serious difficulty of the method lies in determining the "dead space," particularly if adsorption is small.   McBain and Baker weigh the adsorbate in a small bucket containing the adsorbent, suspended in the gas space from a sensitive, calibrated spring of fused silica.

The factors directly influencing adsorption are pressure, temperature, and the nature of adsorbate and adsorbent.

The General Characteristics of Adsorption Equilibrium.

The marked variation in the amount x of a gas adsorbed isothermally by m gm. of adsorbent with the equilibrium pressure p is best expressed in the form of an adsorption isotherm by
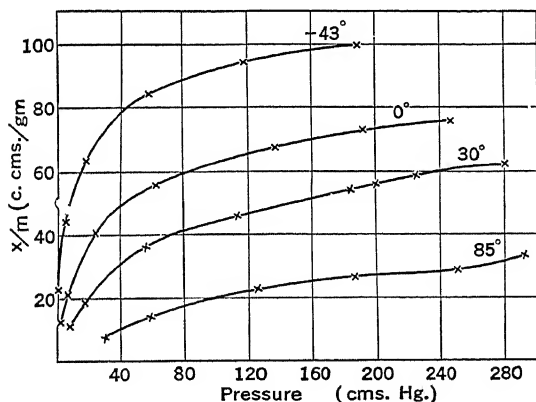
FIG. 2.   The Adsorption of Carbon Dioxide by Charcoal.*

plotting x/m against p.   Typical curves, showing the adsorption of carbon dioxide by charcoal at a number of temperatures, are reproduced in Fig. 2.   The amount of gas adsorbed appears to be a continuous function of the pressure, no breaks being ob-

* RICHARDSON, L. B., *J. Am. Chem. Soc.* 39, 1828 (1917).

served as is the case when definite compounds are formed *
(e.g., the absorption of water by anhydrous copper sulfate).
Two interesting facts appear from the study of such isotherms.
First, adsorption apparently tends to an asymptote, indicating
potential saturation of the surface with the gas; second, at ex-
tremely low pressures the amount adsorbed is directly propor-
tional to the pressure.   The latter is equivalent to stating that
in this range Henry's Law holds for the distribution of the gas
between the two phases.   It has long been recognized that the
empirical equation,†

$$\frac{x}{m} = kp^{\frac{1}{n}} \tag{1}$$

where k and n are constant, can often be used to express the
experimental data on the isotherm.   The simplest technique is
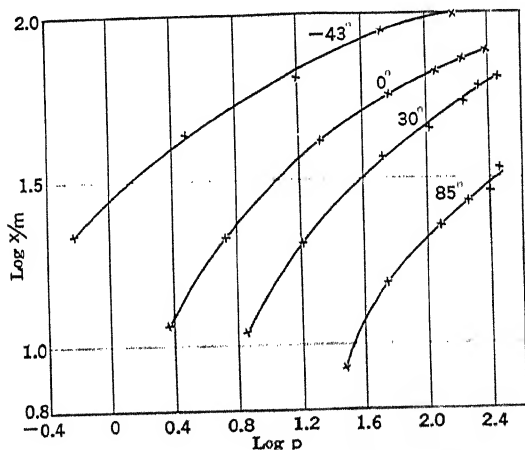
FIG. 3.   Logarithmic Plot of Adsorption of Carbon Dioxide by Charcoal.‡

to plot the logarithm of x/m against the logarithm of p, the
equation requiring that a straight line should result.   However,
the relation always breaks down over sufficiently wide ranges of
pressure, the value of $1/n$ increasing as the pressure is reduced.

* Allmand and his co-workers [Proc. Roy. Soc. **A**, **130**, 610 (1931)], however, have
found that the adsorption isotherms of gases on charcoal consist of a series of very short,
discontinuous, step-wise curves.
† Frequently known as the Freundlich adsorption isotherm.
‡ Richardson, L. B., J, Am, Chem. Soc **39**, 1828 (1917).

This is brought out in Fig. 3, in which the data of Fig. 2 are replotted.   The equation must break down under limiting conditions because it indicates no adsorption maximum at high pressure nor proportionality of adsorption to pressure at low. Its constants have no theoretical significance and are functions of the temperature.   However, for interpolation over moderate ranges of pressure it is frequently useful.

Other than in exceptional cases, adsorption unaccompanied by evidence of chemical action decreases with rise in temperature, although the temperature level required to make it negligible depends on the magnitude of the attractive forces.   With rise in temperature the motion of the gas molecules becomes more rapid, making it more difficult for the attractive forces of the surface molecules of the solid to hold them.   The general variation of the adsorption isotherms with temperature is shown by Fig. 2.   The constant $1/n$ of the Freundlich equation increases * and at the same time the second constant k diminishes.   It is of interest that the general form of the adsorption isotherms of a gas on a given solid remains unchanged above its critical temperature.

In general, the amount of adsorption by a given adsorbent decreases with the volatility of the substance.   Since adsorption is undoubtedly determined by the strength of the intermolecular forces of attraction of the molecules, any physical quantity dependent upon this factor (see Chapter I) would be expected to run parallel to adsorption on a given adsorbent.   Such physical properties are, for example, the van der Waals coefficient a, the boiling point, and the critical temperature.†   Thus Dewar ‡ found that for hydrogen, nitrogen, and oxygen at atmospheric pressure the amounts adsorbed on charcoal at their boiling points were approximately the same, 260 c.cm. per gm.   These rules are by no means without exception, but for a given adsorbent the adsorbabilities of different gases are comparable at corresponding reduced temperatures.   (Table I)

* $1/n$ approaches unity, equivalent to stating that Henry's Law is more nearly obeyed as the temperature rises.

† The boiling point of a liquid in ° K. is approximately two-thirds of its critical temperature, ° K.

‡ DEWAR, J., *Proc. Roy. Soc.* A, **74**. 122, Cambridge University Press (Macmillan) (1904).

Adsorption equilibria are often reversible and approachable from either side. For specific mixtures of gases, the final endpoint has been found independent of the order in which the gases

TABLE I. ADSORPTION OF GASES BY 1 GM. OF CHARCOAL *

Adsorbate	M. Wt.	C.Cm. Adsorbed at 15° C.	B.P. of Gas, °K.	Crit. Temp. of Gas, °K.	van der Waals' Constant, liter² — atm.[b]
SO₂	64	379.7	263	430	3.6
NH₃	17	180.9	240	405	1.8
N₂O	44	54.2	183	309	2.1
C₂H₂	26	48.9	189[a]	308	2.4
CO₂	44	47.6	195[a]	242	2.0
CO	28	9.3	81	134	0.85
N₂	28	8.0	77	126	0.83
H₂	2	4.7	20	33	0.16

[a] Solid phase.
[b] Because van der Waals' equation is only approximate, the values of its constants depend on the physical data actually used in computing them. These are based on critical properties. Values based on heat of vaporization and liquid density at atmospheric boiling point are two or three times as great, although the ratio falls with boiling point.

are adsorbed, but in other cases, notably mixtures of gas and vapor, order of adsorption gives differences in amounts adsorbed which at least persist for long periods.[†]

The quantity of a gas adsorbed varies greatly with the nature of the adsorbent. The more porous a solid, *i.e.*, the greater its surface area for a given mass of material, the greater its adsorptive power (see Table II, but for the behavior of vapors, see p. 80, *et seq.*). Crystalline materials in general have much less capacity as adsorbents than amorphous solids.

Activated carbon is one of the strongest adsorbents known. Charcoal as ordinarily prepared by destructive distillation of wood, bones, sugar, etc., has only a low adsorptive power, but selective oxidation, as in a stream of air at 300 to 400° C., or of steam at 700 to 900°, a process known as activation, greatly increases its adsorptivity. Presumably hydrocarbons or tarry matter, coating the surface of the carbon itself or choking the

* Data from McBain, J. W., *The Sorption of Gases and Vapours by Solids*, Routledge, London (1932).
† Perhaps this is because the rate of approach to equilibrium is greatly reduced. Where the adsorption obtained by gradually raising the pressure is less than that on lowering it, the presence of at least a trace of relatively inert, slightly adsorbed gas in the vessel should be suspected, retarding adsorption rate.

fine pores in the charcoal, are burnt out, while the area of the inner surface and the unsaturation of the surface carbon atoms may also be simultaneously increased.

TABLE II.   APPROXIMATE ADSORPTIONS OF NITROGEN AND CARBON DIOXIDE BY 1 GM. OF DIFFERENT ADSORBENTS AT N.T.P.

Adsorbent	N_2, in c.cm.	CO_2, in c.cm.
Cocoanut charcoal	8.0	90
Graphite	Very small	12
Silica gel	——	11
Alumina	——	5
NaCl crystals	0.0006	0.008
Copper	0.005	0.006
Zeolite (chabasite) 1 c.cm. volume	52.	282

Silica gel is another material with great adsorptive power (for preparation, see pp. 181–186).   It is essentially a hydrated silicon dioxide, in the form of a highly porous solid in which the pores or capillaries are more regular and of larger size than in charcoal.   It differs from the latter in having smaller adsorptive power at very low pressures and in its marked selective adsorptive power for water.   Charcoal at low pressures adsorbs organic compounds more readily than it does water.

An interesting class of adsorbents are the zeolites, hydrous calcium aluminum silicates, in which the water is apparently held very loosely, since no change in their crystalline form occurs on its removal.   When dehydrated, the zeolites * are good adsorbents, the places of the water molecules in the crystal seemingly being taken up by the adsorbed molecules.   Compounds such as ether and benzene are not adsorbed, but others, as ethylene or methyl alcohol, are freely taken up.   Obviously, the zeolite is acting as a molecular sieve, small molecules passing into the interior, where they are distributed throughout the solid, probably at quite definite points in its crystal structure.†   Larger molecules are unable to penetrate the pores of the crystals.

Another important characteristic of adsorption is its heat effect.   Since adsorption decreases with rise in temperature, it

---

* WEIGEL, O., and STEINHOFF, E.. Z. Krist. **61**, 125 (1925).

† To distinguish this from adsorption, McBain suggested the name "persorption."

is accompanied by heat evolution. Any significant comparison must keep in mind the fact that the heat effect per mol of gas adsorbed changes greatly with the amount of gas already on the adsorbing surface. Calling x the total mols of gas adsorbed and Q the corresponding heat effect, it follows that dQ/dx is a function of x/m. This differential coefficient, designated q, decreases as increasing quantities of gas are adsorbed, indicating progressive decrease in the intensity of the adsorptive forces as adsorption progresses; e.g., for ammonia on charcoal, q decreases from 11,000 cals. per gram-mol for the initial amount of gas adsorbed to 7000 cals. for the final. Table III shows the total heats of adsorption Q for various gases on charcoal saturated at atmospheric pressure. The heat of adsorption is considerably larger than the latent heat of condensation of the gas to the liquid phase, more nearly paralleling the latent heat of sublimation, indicating that the attractive forces of the solid for the gaseous molecules are far greater than the mutual attractions of the latter in the liquid state.

TABLE III.   HEATS OF ADSORPTION OF GASES BY CHARCOAL

Gas	Heat of Adsorption gm. cals./mol[a]	Heat of Vaporization[b]	Heat of Sublimation[a]
Argon	3636	1504	4180
Nitrogen	3686	1250	..
Carbon monoxide	3416	1410	3715
Carbon dioxide	7300	2540	6100
Ammonia	7200	5000	7120

[a] GREGG, S. J., The Adsorption of Gases by Solids, Methuen and Co., London (1934).
[b] International Critical Tables, Vol. V, McGraw-Hill.

Chemisorption.

Adsorption relationships are often far more complicated than those outlined above. In the adsorption of oxygen by charcoal at low temperatures (approximately $-150°$ to $-200°$ C.) oxygen is taken up reversibly, i.e., is removable by evacuation. The heat of adsorption is approximately 3700 cals. per mol, the order of magnitude of those in Table III. As room temperatures are approached, the behavior is different, only part of the oxygen can be removed as such by evacuation, and the heat of

initial adsorption is very large, of the order of 70,000 cals. per mol, dropping with increasing adsorption to about 4000 cals. per mol.* At still higher temperatures, the differences are even more marked, practically no oxygen, but rather carbon dioxide and monoxide being obtained on evacuation. Furthermore, at low temperatures the gases evolved are mainly CO_2, the percentage of CO increasing progressively as temperature rises † until nearly pure CO is obtained above 600° (the precise temperature dependent upon the character of the carbon) ‡. Apparently the heat of adsorption rises still further with temperature, probably reaching about 200,000 cals. per mol at 400° C.§ The conclusion seems inescapable that the forces holding the oxygen on the surface of the carbon at high temperatures are fundamentally different in type from those operative at low, and that the former are of the same order of magnitude as primary chemical valence forces. The parallelism of the behavior of oxygen at low temperatures to that of chemically inert gases such as nitrogen, argon, and the like, indicates that here the adsorption is predominantly physical in character, due to the van der Waals forces (pp. 69–70). At high temperatures, the oxygen is

FIG. 4. Oxygen at a Charcoal Surface.‖

evidently held by chemical valence bonds. Figure 4 represents the suggestion of Langmuir as to the nature of the surface of carbon. Atoms in the interior of the solid are undoubtedly held together by primary valence forces and are correspondingly saturated. The surface atoms, on the contrary, while linked by these forces to the interior atoms of the particles, have what are, in essence, free chemical valences available on their outer surfaces for chemical combination with an element such as oxygen. The oxygen-to-carbon chemical valence linkage is so

---

* KEYES, F. G., and MARSHALL, M. J., *J. Am. Chem. Soc.* **49**, 156 (1927).

† It has been claimed that the process can be broken up into two steps.

‡ SHAH, M. S., *J. Chem. Soc.* 2661 (1929).

§ For purposes of comparison, the heat of combustion of solid graphite to give carbon dioxide is 94,300 cals. per mol, that of amorphous carbon about 97,000 cal., and that of the combustion of gaseous carbon has been estimated as 380,000 cals.

‖ LANGMUIR, I., *J. Am. Chem. Soc.* **38**, 2221 (1916).

strong that it is to be expected that large quantities of energy would be released when it is formed and that, once formed, it would break with great difficulty. Indeed, the presence of the oxygen on one side of the carbon atom would undoubtedly greatly reduce the strength of the carbon to carbon linkage on the other side, so that, once thus combined, the carbon atom will be torn out of the surface with the oxygen before the latter can leave alone.

This type of behavior, designated chemisorption, is frequent. Thus, while oxygen adsorbed by silver powder * at room temperature cannot be removed by evacuation, it comes off readily as water vapor on treatment with hydrogen gas at 110° C. Similarly, hydrogen adsorbed by certain metallic oxides (*e.g.*, ZnO, see below) at high temperatures comes off on evacuation or further heating as water and not as hydrogen.

The fact has been emphasized that carbon adsorbs oxygen at low temperatures physically, by a process paralleling liquefaction, in contradistinction to the chemical reaction occurring at high temperature levels. It is well recognized that, except at very low vapor pressures, the rate of liquefaction of a gas is practically instantaneous (provided the heat of condensation can be dissipated sufficiently rapidly), and furthermore this high rate is characteristic of the phenomenon relatively irrespective of the temperature level. It is not surprising to find that the same is true of the process of physical adsorption. It is equally well recognized that, in sharp contradistinction to physical adsorption, chemical reactions in general are characterized by extraordinarily high change of rate with the temperature, the rates falling off at sufficiently low temperatures to values so small that they cannot be measured. If one plots the quantity of hydrogen adsorbed on zinc oxide against the temperature, one obtains the curve shown in Fig. 5. In a relatively narrow temperature range there is actually an increase in amount of adsorption. That the adsorption below this temperature range is physical is shown by the possibility of removing the adsorbed gas by evacuation—a conclusion also confirmed by the low value of the heat of adsorption, approximately 1900 cals. per mol at

* BENTON, A. F., and ELGIN, J. C., *J. Am. Chem. Soc.* **48**, 3027 (1926).

0° C. On the other hand, hydrogen absorbed at high tempera-
ture can be removed only as water and its heat of adsorption
is 20,000 cals. (between 300 and 444° C.). Finally, the rate of
adsorption at the high temperature levels changes rapidly with
the temperature, exhibiting in
this respect the characteristics
of a chemical reaction, but
changes slowly in the portion
of the curve of Fig. 5 between
about 250 and 450° K., in which
total adsorption is rising with
temperature. In contradis-
tinction, the rate at low tem-
perature levels is high and
changes but slowly with the
temperature. Similar relations
have been found in the adsorption of hydrogen on nickel. Thus,
adsorption rates offer additional confirmation of the chemical
character of high temperature adsorption in these cases.
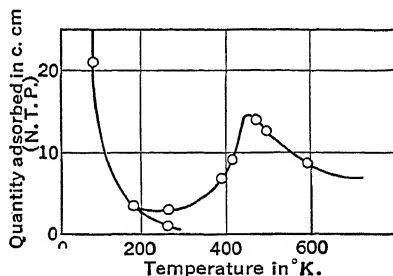
FIG. 5. Adsorption of Hydrogen
on 20.0 gms. Zinc Oxide at 400 mm.
pressure.*

The consensus is swinging to an interpretation of the influ-
ence of temperature on chemical reaction rate as due to an
activation of one or more of the reacting molecules by an increase
of thermal energy to a value so great that, on collision of the
proper type, molecular equilibrium is upset and reaction occurs.
The momentary molecular energy or activation energy neces-
sary for this result varies from one reaction to another and is
given by the equation,

$$\frac{d \ln K}{dT} = \frac{E}{RT^2} \tag{2}$$

where K is the velocity coefficient at absolute temperature T,
E the activation energy, and R the gas constant. In conse-
quence, gas adsorption of this type has been called activated
adsorption. The evidence indicates that there may be a definite
and important relation between the activity of the adsorbed gas
in the sense just described and its capacity for chemical reaction
with other molecular species which may be in the neighborhood,
i.e., with the catalytic capacity of the solid adsorbent.

* TAYLOR, H. S. and SICKMAN, D. V., J. Am. Chem. Soc. 54, 602 (1932).

The Adsorption of Vapors.

As the temperature is lowered the adsorption of a given gas by any specific agent increases. The shape of the adsorption isotherms most frequently encountered is that indicated diagrammatically in Fig. 6. In this figure, the pressure is plotted, not in normal units, but as a fraction of the saturation pressure of the pure liquid adsorbate at the temperature in question (the curves therefore being restricted to adsorbate temperatures below the critical).
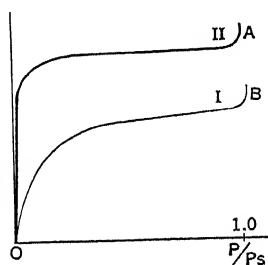
FIG. 6. Adsorption of a Vapor on a Porous Adsorbent.

In general, the lower the reduced temperature * (T_R) of the adsorbate vapor, the sharper is the bend in the curve, i.e., drop in temperature causes a change from I to II. It is important to emphasize how steep the left end portion of curve II may be; thus, there are activated charcoals capable of picking up 85% of their weight of chlorpicrin at a temperature at which that material has a vapor pressure of approximately 15 mm., and yet the vapor pressure of the adsorbed gas is undetectable.† The apparent asymptote indicated by the general trend of the curve rises somewhat as temperature decreases.

In many cases, there develops a sudden rise at the end of the curves shown in Fig. 6(A, B) as the saturation pressure of the gas is approached at any given temperature. It is probably due to bulk condensation of vapor in pores of the adsorbent (see p. 86). There is doubt as to whether the rise is abrupt and discontinuous, or a sharp bend in the curve. For a given porous adsorbent approximately equal quantities of all vapors are adsorbed at saturation, provided they are expressed as liquid volumes. Thus Coolidge,‡ using a specific activated charcoal at 0° C., found the maximum volumes adsorbed varied only from 0.424 ml. per gm. for water to 0.490 ml. for ether. Other vapors

* Reduced temperature is defined as the temperature divided by the critical temperature of the material in question, both expressed on the absolute scale.

† This fact, together with the remarkable universality of the adsorptive power of charcoal properly prepared, is the basis of its use as a protective agent in gas warfare.

‡ COOLIDGE, A. S., *J. Am. Chem. Soc.* **46**, 596 (1924).

investigated were carbon tetrachloride (0.435 ml. per gm.), chloroform (0.442), carbon disulfide (0.450), methyl alcohol (0.450), ethyl formate (0.454), benzene (0.467), and methyl acetate (0.490).

Study of the data indicates, however, that Fig. 6 is not truly representative for all materials. Thus, Fig. 7 shows the approximate shape of the curves for water and benzene on both silica gel and charcoal. The differences between a number of liquids on the same charcoal are shown in Fig. 8, due to Coolidge, where the logarithm of the amount adsorbed, expressed as a fraction of the quantity adsorbed
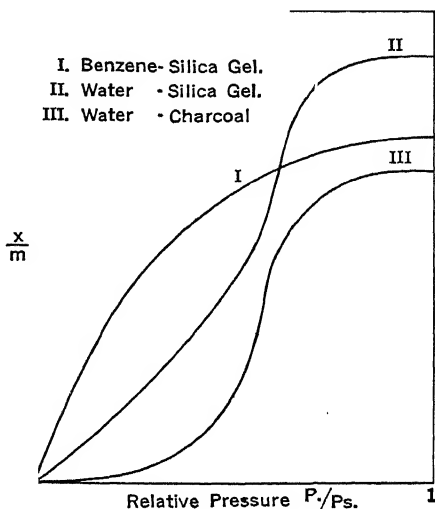
FIG. 7. Adsorption of Vapors on Silica Gel and Charcoal.

at saturation, is plotted against the logarithm of the pressure expressed as a fraction of the vapor pressure of the pure liquid at the same temperature. It seems clear that two types of isotherms may be distinguished, one belonging to non-polar substances, such as ether, benzene, etc., regular throughout its course, while the other is exhibited by polar compounds such as water, methyl alcohol, and ammonia. At low pressures the latter isotherms are convex towards the axis of pressures. The diagram emphasizes the inadequacy of the Freundlich equation,
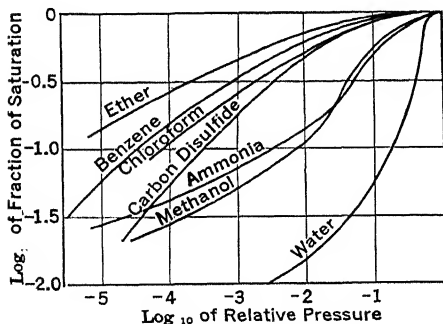
FIG. 8. The Adsorption of Vapors on Charcoal.*

which requires straight lines for each substance.

* Data of COOLIDGE, A. S., *J. Am. Chem. Soc.* 46, 596 (1924).

Theories of Physical Adsorption.

Data such as just outlined have led to the assumption of two major mechanisms of physical adsorption. The most commonly met is probably so-called physical or van der Waals adsorption, attributed to the stray fields of force of the surface molecules of the adsorbent. The second, capillary adsorption, may result at high relative pressures on porous solids from condensation occurring in the pores.* The complications of the phenomena often make difficult the disentanglement of the mechanisms.

Langmuir attempted to minimize disturbing factors by a study of plane surfaces. The adsorption per unit surface area of various gases on mica, glass, and platinum was measured at low pressures, large areas being secured by arranging sheets of the solid in layers. It was found that the quantity of gas adsorbed corresponded to gas layers of one, or at most, only a few molecules thick. The curve for the adsorption of nitrogen on mica at 90° K. is shown in Fig. 9. Two points are noticeable: at low pressures the data show a nearly constant slope, while at high pressures the curve tends to an asymptotic, saturation value. This asymptote corresponds to less than a monomolecular layer of nitrogen, only about 20% of the total available surface being covered with adsorbed molecules. On mica and glass similar results were obtained for the other gases examined. In all cases the order of adsorption was methane, carbon monoxide, nitrogen, argon, and oxygen. The fact that adsorption varies greatly from gas to gas and is so incomplete suggests that it occurs only at spots on the surface where there are particularly strong fields of force.
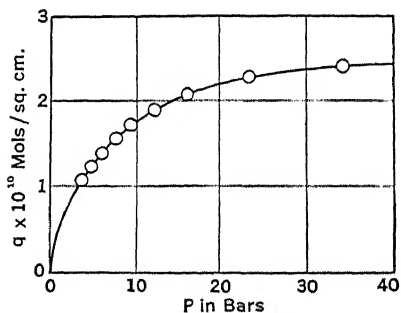
FIG. 9. The Adsorption of Nitrogen on Mica at 90° K.†

Langmuir ‡ was able to give a satisfactory explanation of these and similar experimental results by treating adsorption as

---

* Possibility of diffusion into the interior of the solid must be borne in mind.
† Data from LANGMUIR, I., *J. Am. Chem. Soc.* **40**, 1361 (1918).
‡ LANGMUIR, I., *J. Am. Chem. Soc.* **40**, 1361 (1918).

an effect due to the time lag between a molecule hitting a surface and rebounding from it.   When a gas molecule impinges upon a solid surface it does not, in general, rebound elastically but is held for an appreciable time close to the surface by its field of force.   The gas molecules may, therefore, be considered as having condensed on the surface, from which thermal agitation may subsequently cause them to break away or to evaporate. Adsorption is regarded by Langmuir as the direct consequence of this time lag between condensation and evaporation, its magnitude depending on the balance between the surface forces attracting the gas molecules and the disgregating tendencies, particularly thermal agitation.   Further, since the effectiveness of the field of force of the surface is extremely limited in extent (cf. p. 5, *et seq.*), comparable to a molecular diameter, in absence of large mutual attractions of the adsorbed molecules, a limiting value of the adsorption should be reached when the surface is covered with a monomolecular layer.   Certain atoms or molecules in a crystal surface will probably have stronger external fields of force than others; hence, adsorption will tend to occur first at these points, the weaker atoms or molecules becoming effective only at higher concentrations.   The points on the surface effective in adsorption have been termed by Langmuir "elementary spaces."   In many cases a surface may be made up of different types of elementary spaces, each with a different adsorptivity.   In others, the molecules of adsorbate may be so large that they cannot occupy adjacent elementary spaces on the crystal, giving rise to steric hindrance.   These factors still further complicate the adsorption process.

Langmuir considers first the case of a surface having only one kind of elementary space, capable of holding only one adsorbed molecule per space, and assumes that there is no interaction between adsorbed molecules.   Let μ represent the number of mols of gas striking one square centimeter of the surface per second.   Of this amount, a certain fraction α will condense and be held by the surface forces until subsequent evaporation. The rate at which the gas condenses on the bare surface is, therefore, $\alpha\mu$.   The maximum number of molecules which can be adsorbed on the surface is limited by the number of ele-

mentary spaces, N_0, per sq. cm. of surface. For very low gas concentrations the force acting between the first adsorbed gas layer and any second layer that would form upon it, will be very small. Thus, any molecule condensing upon the first layer will evaporate almost instantaneously, giving, in effect, a reflection. It follows that if θ is the fraction of surface which is covered,

$$\text{rate of evaporation} = v\theta,$$
$$\text{rate of condensation} = \alpha\mu\,(1 - \theta),$$

where v is the rate at which gas molecules would evaporate from the surface were it entirely covered.

At equilibrium the rates of condensation and evaporation must be equal:

$$\alpha\mu\,(1 - \theta) = v\theta,$$

whence

$$\theta = \frac{\alpha\mu}{v + \alpha\mu}. \tag{3}$$

If the constants of (3) are combined by making

$$\sigma = \frac{\alpha}{v},$$

it may be written

$$\theta = \frac{\sigma\mu}{1 + \sigma\mu}.$$

Calling q the number of mols of gas adsorbed per sq. cm. and N the Avogadro constant,

$$\frac{Nq}{N_0} = \theta = \frac{\sigma\mu}{1 + \sigma\mu}. \tag{4}$$

The rate at which the gas molecules hit the surface can be calculated from the kinetic theory, by the Herz-Knudsen equation,

$$m = \frac{p}{\sqrt{2\pi MRT}} = kp \tag{5}$$

where m is the number of gram-molecules hitting 1 sq. cm. of the surface per second, p the pressure in dynes per sq. cm., M the molecular weight, T the absolute temperature, and R the gas

constant.   Combining (4) and (5) it follows that

$$q = \frac{abp}{1 + ap} \tag{6}$$

where b is N_0/N and the constant a is a function of the temperature and the gas under consideration, being proportional to m and inversely proportional to v.   The relation is known as the Langmuir adsorption isotherm.

At very low pressures the relation reduces to $q = abp$.   Thus, as found by Langmuir, the adsorption is directly proportional to the pressure.   Since it is inversely proportional to v, it increases rapidly with decreasing temperature.   At very high pressures the isotherm reduces to $q = b = N_0/N$.   Hence, the asymptotic adsorption should be substantially independent of temperature, except for minor variation of N_0 with temperature. This independence of temperature is not found experimentally and has been the basis of criticism of the equation.   Otherwise, Langmuir found good agreement between his results and the equation, as shown by Table IV.   While the assumption of only

TABLE IV.   ADSORPTION OF NITROGEN ON MICA
(5750 SQ. CM. IN AREA) AT 90° K.*

P, in bars [a]	q, in cu. mm. (N.T.P.)	$q \times 10^{10}$, mols per sq. cm.	$q \times 10^{10}$ (calculated)
34.0	33	2.39	2.39
23.8	30.8	2.235	2.24
17.3	28.2	2.045	2.07
13.0	25.5	1.85	1.91
9.5	23.9	1.73	1.705
7.4	21.6	1.565	1.53
6.1	19.0	1.38	1.39
5.0	17.0	1.23	1.25
4.0	15.1	1.095	1.10
3.4	13.4	0.97	0.99
2.8	12.0	0.87	0.87

[a] Dynes per sq. cm.

one kind of elementary space on which the isotherm was derived is probably exceptional, it is found to hold in many cases, particularly if the pressure range is not excessive.   McBain and

* LANGMUIR, I., *J. Am. Chem. Soc.* **40**, 1361 (1918).

Britton,* who determined the sorption of nitrogen, nitrous oxide, and ethylene on charcoal over a range of pressures up to 60 atmospheres, also found that their results could be satisfactorily expressed by the use of the Langmuir equation.

If an adsorbent has more than one kind of elementary space, each can be regarded as acting independently, so that an equation is obtained which is the sum of a number of terms of the type on the right of equation (6). Thus, if the several kinds of elementary spaces occupy fractions β_1, β_2, etc., of the surface, we have:

$$q = \frac{\beta_1 \sigma_1 \mu}{1 + \sigma_1 \mu} + \frac{\beta_2 \sigma_2 \mu}{1 + \sigma_2 \mu} + \quad \text{etc.} \tag{7}$$

The difficulty of applying and using a relation of this sort is obvious.

The Langmuir theory attributes adsorption essentially to the formation on the surface of the adsorbent of a layer of gas one or at most two to three molecules thick. Eucken and Polanyi, among others, attacked the problem from a different angle, postulating the formation of a thick film. It is assumed that the actual force between the solid surface and a molecule in the gas space is a function only of its distance from the surface, independent of the temperature and concentration of the gas molecules. The work ϵ required to bring a mol of the gas from infinity to a definite distance from the surface is known as the "adsorption potential," and will be $\int v \cdot dp$, or for a vapor following the gas laws $RT \ln p/p_s$ where p_s is the saturation pressure. This quantity is a function of the volume adsorbed ϕ, $\epsilon = f(\phi) = f(x/\rho)$ where ρ is the density of the adsorbate, assumed to be liquid. From a single isotherm, one can plot the adsorption potential ϵ against the volume adsorbed. Because this should be independent of temperature, one can now compute the isotherms for different temperatures, providing the change of density with temperature is known. In many cases the Polanyi theory is confirmed in remarkable degree, and indicates that the surface forces are in fact relatively independent of temperature. Unfortunately, the functions differ for various adsorbates, and utility of the method is limited.

It has already been mentioned that adsorption at high pressures is often greater than can be accounted for by the lower portion of the adsorption curve and that this is explained by the occurrence of bulk condensation of liquid as such. Zsigmondy, Andersen, Patrick, and others suggested that adsorption of gases by porous solids may be due to the fact that the vapor pressure of the curved surface of a liquid in a capillary is less than that

* McBain, J. W., and Britton, G. T., *J. Am. Chem. Soc.* 52, 2198 (1930).

of a plane surface and that porous adsorbents, such as charcoal and silica gel, are made up of a large number of extremely small capillaries of varying size. The relation between the radius r of the curved surface of a liquid and its vapor pressure, first derived by Rayleigh, is

$$r = \frac{2\sigma M}{\rho RT \ln \dfrac{p}{p_s}} \qquad (8)$$

where σ is the surface tension, ρ the density and M the molecular weight of the liquid, p the vapor pressure of the liquid at the curved surface, p_s the saturation pressure of the plane liquid surface at the same absolute temperature T, and R the gas constant.* The equation indicates not only that small drops have excess vapor pressure, but also that vapor will tend to condense first in the smaller capillaries of an adsorbent and only later in the larger ones. However, calculating the necessary size of capillary from adsorption data shows that for pressures much below saturation, the radii thus computed are of molecular dimensions or smaller (Table V). Capillary condensation, while

TABLE V. ADSORPTION OF SULFUR DIOXIDE ON SILICA GEL AT —34° C.†

p, mm.	x/m, c.cm. (N.T.P.)	$\log p/p_s$	$r \times 10^{-8}$, cm.	Capillary diameter / Molecular diameter
4.22	53.31	1.741	3.75	1.8
13.85	81.5	1.224	5.34	2.6
29.02	108.18	0.904	7.23	3.5
50.07	137.17	0.666	9.80	4.8
73.45	162.6	0.500	13.05	9.5
116.62	192.1	0.300	21.8	9.6
205.59	206.5	0.053	123	60
232.66	260.0	0		

no doubt often occurring in the high pressure range of adsorption, fails to account for the high adsorption frequently occurring

* A liquid can be transferred by distillation isothermally and reversibly from one spherical drop to another. For such a process the net work, i.e., the free energy change, must be zero. Equating the distillation work to the change in surface energy results in this formula.

† Data from MacGavack, J., Jr., and Patrick, W. A., J. Am. Chem. Soc. 42, 946 (1920).

at very low relative pressures.   Furthermore, it cannot account for specificity of adsorption or for adsorption at planar surfaces.

Adsorption and Catalysis.

There can be no doubt that adsorption plays an important part in contact catalysis, *e.g.*, the oxidation of sulfur dioxide or ammonia on platinum or the hydrogenation of organic compounds by nickel.   To Faraday and other early investigators the mechanism of solid catalysis remained a mystery, but the concept of monomolecular adsorption throws considerable light on the process.   An illustration is given by Langmuir's * explanation of the catalytic oxidation of carbon monoxide by platinum.   At low temperatures (200 to 400° C.) the rate of reaction is proportional to the partial pressure of oxygen and inversely proportional to the pressure of carbon monoxide, the reaction rate also increasing rapidly with rise in temperature.   At higher temperatures (500–800° C.) the reaction rate is independent of temperature, is proportional to the pressure of carbon monoxide when oxygen is in excess, and proportional to the pressure of oxygen when carbon monoxide is in excess.   It is also found that, while carbon dioxide is little adsorbed by platinum, carbon monoxide is strongly adsorbed at low temperatures, but oxygen is very strongly adsorbed even at high temperatures.   It follows
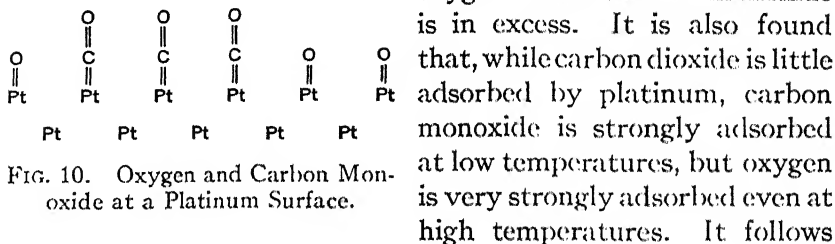
Fig. 10.   Oxygen and Carbon Monoxide at a Platinum Surface.

that at low temperatures carbon monoxide and oxygen will both be adsorbed, competing for the empty elementary spaces on the platinum and producing a surface of the type in Fig. 10.   If it now be assumed that carbon monoxide molecules striking adsorbed oxygen molecules can react, granting adequate impact energy, while oxygen molecules striking carbon monoxide molecules do not react, since they strike the oxygen atom with its completely satisfied valences, it is apparent why carbon monoxide acts as a poison.   It takes up spaces on the surface which might have been occupied by oxygen molecules capable of reac-

* LANGMUIR, I., *Trans. Far. Soc.* 17, 621 (1922).

tion. The adsorption of carbon monoxide decreases with temperature, so that the rate of reaction will increase. However, at about 450° C. the adsorption of carbon monoxide becomes so small that the percentage of the surface occupied by it is negligible, so that the rate of reaction at constant oxygen pressure becomes substantially independent of temperature. At the same time, if carbon monoxide is present in large excess, every oxygen atom adsorbed will react almost immediately. Hence, the reaction rate will be proportional to the rate of adsorption of oxygen, and this is in turn proportional to its partial pressure. Similarly, when oxygen is in excess, there will always be sufficient oxygen molecules adsorbed for the rate of reaction to be governed by the partial pressure of the carbon monoxide.

It must be recognized (p. 34) that no surface, even of a single crystal, is completely homogeneous. There is no doubt that this heterogeneity affects both adsorption and catalysis, and in differing degree. Thus, small traces of an impurity, sufficient to poison the material as a catalyst, may be insufficient greatly to affect its adsorptive power. In illustration, Pease and Stewart * found that a trace of carbon monoxide is sufficient to reduce the catalytic effect of a sample of reduced copper on the combination of ethylene and hydrogen to 15% of its original value, despite continued adsorption of the reacting gases. Again, if a colloidal platinum catalyst is progressively poisoned, its activity towards the hydrogenation of organic compounds is reduced step by step; e.g., with small amounts of carbon disulfide the hydrogenation of propyl ketone is prevented, larger amounts stop the reduction of piperonal, but still further amounts are required to prevent the reduction of nitrobenzene. Such facts lead inescapably to the conclusion that different regions of a surface may have different activities; those of high activity are the so-called active centers. Thus, according to Taylor †, on the surface of a nickel catalyst the nickel atoms at peaks and along cracks may be more unsaturated and reactive than the others.

It is found that sometimes the addition to a catalyst of materials, themselves catalytically inactive, greatly increases

* PEASE, R. N., and STEWART, L., J. Am. Chem. Soc. 47, 1235 (1925).
† TAYLOR, H. S., Proc. Roy. Soc. A, 108, 105 (1925),

activity. Such substances are termed promoters. The general mechanism of promoter action is uncertain; one obvious explanation is that the inert material breaks up the surface of the catalyst, thus multiplying active centers; or the promoter may serve as adsorbent for one of the reacting gases unadsorbed by the catalyst itself.

Adsorption from Solution.

Phenomena similar to gaseous adsorption occur at solid-liquid interfaces. A typical example is the decolorization of an aqueous dyestuff solution by boiling with animal charcoal. On washing the charcoal with alcohol or other suitable solvent, the dyestuff may be extracted. This indicates that physical adsorption rather than chemical combination has occurred and emphasizes the great influence of the solvent on the adsorption of a particular substance. The solute molecules do not behave merely as gas molecules, with the solvent as an inert diluent.

For dilute solutions at a given temperature, Freundlich showed the validity of an equation analogous to that for gases,

$$\frac{x}{m} = kc^{\frac{1}{n}} \qquad (9)$$

where k and n are constants for a particular system. Here too the values of the constants have no theoretical significance and vary widely: for most systems $1/n$ lies between 0.1 and 0.7. The Langmuir equation,

$$\frac{x}{m} = \frac{abc}{1 + ac} \qquad (10)$$

can also be used for adsorption from solution. Table VI shows results for the adsorption of the fatty acids on charcoal correlated by both equations. The Langmuir equation is less satisfactory, but its constants possess certain physical significance. The constant b is particularly important, since it constitutes the theoretical limit of the adsorption; from it one may estimate the number of molecules held on the saturated surface of the adsorbent.

For dilute solutions the equations of Freundlich and Langmuir are satisfactory, but as concentration increases the adsorp

TABLE VI. ADSORPTION BY NORITE FROM AQUEOUS ORGANIC ACIDS (0 TO 0.2M) *

Acid	k	$1/n$	Average % Error Between Observed and Calculated Values (Freundlich)	a	b	Average % Error Between Observed and Calculated Values (Langmuir)
Formic	2.47	0.435	2.6	0.159	1.710	1.0
Acetic	2.46	0.351	2.6	0.266	1.736	7.0
Propionic	2.46	0.236	1.8	0.491	1.885	2.0
n-Butyric	2.46	0.177	2.3	0.863	1.957	2.0
n-Valeric	2.84	0.182	2.8	0.872	2.154	5.8
Caproic	3.03	0.175	1.8	4.636	1.892	5.4

tion becomes less than that indicated by either, and may even in some cases become negative.†

As in the case of gas adsorbents, the adsorbing powers of different solids vary enormously. Roughly, adsorbents can be grouped in two classes. Hydrophilic adsorbents, such as silica gel and all the oxide gels (p. 181), including kieselguhr, have marked affinity for water and polar compounds, adsorb strongly from organic solvents, but have little adsorptive power from aqueous solutions. On the other hand, hydrophobic adsorbents, of which charcoal is the best example, adsorb strongly from water, for which they have little affinity.

A given solute tends to be adsorbed by a given solid most strongly from solvents in which it is least soluble. For solvents of a given polarity, the concentrations of a given solute on the adsorbent, $(x/m)_0$, when in equilibrium with a saturated solution, are approximately the same.‡ Calling the solubilities in the various solvents s_1, s_2, etc., and the corresponding constants of the Freundlich equation k_1, k_2, n_1, n_2, etc., $(x/m)_0 = k_1 s_1^{1/n_1} = k_2 s_2^{1/n_2}$, etc., Table VII shows that for liquids of similar polarity the deviations are not great.

* Data of LINNER, E. R., and GORTNER, R. A., J. Phys. Chem. 39, 35 (1935) (The Williams and Wilkins Co.).

† Usually one computes adsorption by subtracting the solute corresponding to the imbibed solvent from the total solute picked up by the solid. If the solvent be more highly adsorbed than the solute, so-called negative adsorption is observed, although true adsorption may still be considerable. Obviously the method is correct only when no solvent is adsorbed, a condition probably never fulfilled in practice.

‡ They would, of course, be identical except for competitive adsorption of solvent.

Few rules can be drawn for predicting adsorption from solution. There appears to be no general connection between the adsorption from a given solvent and its surface tension, although

TABLE VII.   ADSORPTION OF PICRIC ACID AT 20° C. ON
CHARCOAL *

Solvent	$1/n$	k	Solubility	$(x/m)_0$
Water	0.17	1.53	1.11	0.155
Ethyl alcohol	0.24	0.537	6.06	0.827
Acetone	0.28	0.288	54.59	0.882
Chloroform	0.27	0.280	2.29	0.350
Benzene	0.19	0.358	9.56	0.550
Toluene	0.13	0.417	9.79	0.561
Ethylbenzene	0.26	0.360	4.82	0.542
Isopropylbenzene	0.16	0.285	10.97	0.418
Nitrobenzene	0.15	0.199	25.25	0.323

adsorption often parallels it.   Similarly, no connection has been traced between the lowering of the surface tension of a solvent by the adsorbed substance at the solvent-vapor interface and the amount of adsorption.†

Commercial Applications of Adsorption.

In Table VIII are given certain adsorbents in general industrial use.

TABLE VIII.   ADSORBENTS IN INDUSTRIAL USE

Adsorbent	Industrial Use
Charcoal (in various forms, activated by different methods)	Gas masks, solvent recovery, sugar refining, removal of industrial odors, decolorizing of many liquids, etc.
Bone char (animal charcoal)	Sugar refining
Fuller's earth	Petroleum and vegetable-oil refining
Bentonites,[a] etc.	
Silica gel	Dehydration and purification of gases

[a] Acid-treated clays are frequently sold under trade names.

The adsorbents of industrial importance for gases are charcoal and silica gel.   A number of activated carbons, differing in preparation and adsorptive power, are on the market.   In gen-

* SATA, N., *Kolloid Z.* **49**, 275 (1929).
† Obviously, however, this is no indication of the change in interfacial tension at the solid-solvent interface, which is the important factor in the Gibbs adsorption isotherm.

eral, the adsorbent in granular form is packed in towers and the gas containing the adsorbate passed through it until the adsorbent is nearly saturated. The gas stream is then deflected, usually into a tower of fresh char, and the adsorbent stripped ot the adsorbate, usually by steam distillation. Thus, benzol and most volatile organic liquids can be removed from charcoal by superheated steam, subsequent cooling rendering the charcoal again ready for use. Active charcoal is particularly valuable in the recovery of solvent vapors which are very largely diluted with air, as from the drying of rubber cements, dry-cleaning solvents, lacquers, etc.*

For solutions, adsorbents are employed in two ways, namely, by percolation and contact. In the former, the adsorbent, which should be granular and retain its shape, is packed in towers and the solution to be purified allowed to percolate through it. In the contact method, the finely divided adsorbent is agitated with the liquid to be treated, and the adsorbent removed by settling, filtration, or a combination of both.

One of the important uses of adsorbents is as a decolorant in sugar refining. Bone charcoal, through which the sugar solution is allowed to percolate, has been in use for this purpose for over a century. It differs from ordinary activated charcoal in that it is prepared by the ignition of bones and contains about 75% of tricalcium phosphate. This inorganic material serves as a skeleton, full of minute pores and channels, over which the carbon is distributed, thus forming a highly porous structure with large surface and high adsorptivity. About one pound of bone char is required for each pound of raw sugar, though this naturally varies widely with the type of sugar being treated. The char is placed in large vertical cylinders which have porous false bottoms. The highest grade sugar solutions at 160 to 170° F. are sent through the fresh char, followed by lower grades until the char no longer decolorizes efficiently. The char is washed with water and then revivified by heating with a limited supply of air in such a way that the adsorbed impurities are selectively oxidized. A char may be revivified up to 200 times before it becomes necessary to discard it. The amount of ad-

* The use of charcoal in gas masks requires no discussion.

sorption as measured at equilibrium approximates the Freundlich equation, but many factors have to be considered in the commercial process. The size of the particles exerts an effect, rate of decolorization being more rapid the finer the char (Fig. 11). Decolorization is also greater the higher the temperature and the larger the quantity of char used. Low pH favors adsorption, but causes sugar losses due to inversion. The effectiveness of a given char varies greatly with the type of color to be removed, *i.e.*, of the sugar solution to be decolorized.
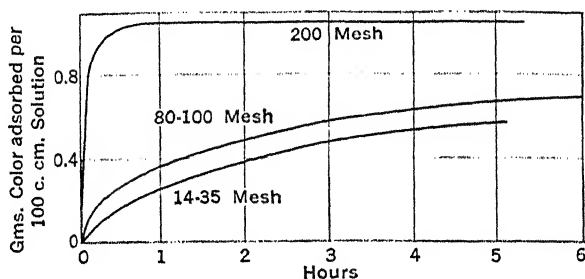
FIG. 11. Decolorization of Sugar Solution by Charcoal.* Size of sample 100 c.cm.; Wt. of color (caramel) per 100 c.cm., 1.1 gm.; Wt. of char. per 100 c.cm., 10 gm.

Certain activated charcoals used in the contact process are much more active as decolorizing agents than bone char, leading to economies in capital cost of equipment, amounts of wash water to be treated, etc. On the other hand, most of the present refineries in the United States have been built for the bone char process and it is claimed that bone char is more effective in adsorbing inorganic materials, which interfere with the crystallizability of the sugar and hence with yield.

Activated carbons are widely used in the purification of liquids and solutions, as in decolorizing synthetic organic chemicals.

In the decolorization of petroleum oils, clay or fuller's earth is used both by percolation of the oil through a relatively thick bed of coarse clay particles and by contacting the fine clay in the hot oil. Percolation is always carried out at somewhat

* FREIHEIT, A. J., *Mass. Inst. Tech. Thesis* (1924).

elevated temperatures, the oil being frequently diluted with naphtha or an equivalent solvent to lower its viscosity. Contacting is usually conducted at a temperature level considerably higher, sometimes up to 500 to 600° F. The clay is often activated by treatment with acid before use (p. 452). At high temperatures the clay exerts a definite catalytic cracking effect, but at lower temperatures both contacting and percolation are evidently predominantly adsorptive. However, the adsorbed materials, color bodies, etc., are often firmly held and can be removed only by ignition. This revivification by burning should be conducted at as low temperature levels as practicable and, while it results in lost adsorption efficiency, none the less the clay can often be used many times. Revivification always results in loss by dusting. Contacting clays are rarely revivified. The organic matter need not be completely removed in the revivification, but must be thoroughly carbonized. Where contacting is conducted under conditions involving cracking, the volatile constituents must be removed by suitable subsequent distillatory operations, *e.g.*, steam stripping in vacuum. The tenacity with which the adsorbed materials are held by the clay may perhaps be due to catalytic polymerization by the clay.

BIBLIOGRAPHY

FARADAY SOCIETY DISCUSSION, *The Adsorption of Gases by Solids*, Jan., 1932.

FREUNDLICH, H., *Colloid and Capillary Chemistry*, translated by H. S. Hatfield, Methuen and Co., London (1926).

GREGG, S. J., *The Adsorption of Gases by Solids*, Methuen and Co., London (1934).

HÜCKEL, E., *Adsorption and Kapillarkondensation*, Akad. Verlag. m.b.h., Leipzig (1928).

McBAIN, J. W., *The Sorption of Gases and Vapours by Solids*, Routledge, London (1932).

THOMAS, A. W., *Colloid Chemistry*, McGraw-Hill (1934).

Chapter VI

Suspensoids

If a suitable solute, such as sugar, be brought into contact with water it disperses in the liquid, forming a physically homogeneous solution, the properties of which indicate that the sugar is broken down in the liquid into its ultimate molecules. If finely divided quartz be brought in contact with water, it too can be made to disperse in the liquid. In this case, however, unlike the sugar, there is no evidence whatever to indicate a breakdown or further dispersion of the quartz particles after coming into contact with the liquid. Furthermore, the quartz particles in suspension retain essentially all their original physical properties unchanged, whereas the dispersed sugar particles possess properties differing profoundly from those of the original solid, *e.g.*, greatly increased capacity for diffusion through membranes of low permeability.* By proper choice of the liquid dispersing medium and the material dispersed, one can find all degrees of dispersion, between that of an electrolyte solution at the one extreme and that of a relatively coarse suspended solid at the other. Furthermore, study of these systems indicates that a most important property in which they differ is the size of the ultimate particle of the dispersed material. Where the particle size is of the order of magnitude of ordinary molecules, the dispersion is called a solution. Where the solid exists in the form of relatively coarse suspended particles, the system is called a suspension. It is found experimentally that, as particle size decreases, the stability of the system tends to increase very greatly. Dispersions in which the particles are large compared with the magnitude of ordinary molecules but still small enough to possess at least reasonable stability are in general called

---

* Perhaps the most striking change in properties characteristic of dispersion in a liquid are those shown by electrolytes.

colloidal dispersions, colloidal solutions, or sols; if the properties indicate that the particles are relatively inert suspended solids, the dispersion is described as a suspensoid, to distinguish it from the less stable suspensions. There are no sharp lines of demarkation between the three types of dispersion. With change in particle size, properties change progressively without discontinuity. The range of colloidal solutions may be defined very roughly as involving dispersed particles of dimensions from 1 to 500 mμ.*

Theoretically, since matter can exist in three forms, liquid, solid, and gaseous, a number of types of disperse systems are possible, as shown in Table I. In practice, solid-liquid and liquid-liquid dispersions are of greatest industrial importance.

TABLE I.   DISPERSE SYSTEMS

Disperse Phase	Dispersion Medium	Example
Gas	Gas	None [a]
Liquid	Gas	Mist
Solid	Gas	Smokes
Gas	Liquid	Foams
Liquid	Liquid	Emulsions (cream)
Solid	Liquid	Sols (gold sol)
Gas	Solid	Meerschaum
Liquid	Solid	[b]
Solid	Solid	Ruby glass

[a] Gases are completely miscible; hence colloidal dispersions of one gas in another are impossible.
[b] Quickly frozen sea-water raised in temperature to just above its eutectic point.

There are two important problems regarding suspensions which must first be studied. Granting that the suspended particles are heavier than the liquid around them, one would expect the particles to settle out completely. It is found experimentally that in many cases, especially where the particles are small in size, there is relatively little settling, sedimentation never progressing to the point of complete precipitation of the suspension. Thus, certain of Faraday's original gold dispersions in water, prepared in 1857, are still preserved at the Royal Institution and have deposited relatively little of their gold content. In the second place, whereas in many suspensions the

* The unit μ is 10^{-3} mm. $= 10^{-4}$ cm.   The unit mμ is $10^{-3} \mu$.   The so-called Ångström unit Å is $10^{-4} \mu$ or 10^{-8} cm.

particles agglomerate into larger clumps more or less rapidly, in many others no such agglomeration occurs. Both these phenomena, which so profoundly influence the stability of suspensions, demand explanation.

Stokes' Law.

An insoluble particle of a substance placed in a liquid tends to sink or float, depending upon the relative densities of the solid and liquid. If the particle is small, its velocity of fall will rapidly become constant at a low value. Inspection of the liquid in its immediate neighborhood shows that it flows around the particle in a characteristic streamline fashion, termed viscous flow.* Experimentally it is found that the frictional force tending to resist the motion of the particle is proportional to its velocity. By hydrodynamic analysis Stokes showed that, for a spherical particle moving in viscous motion in a fluid medium of infinite extent, the frictional force F, opposing its motion, is given by

$$F = 6\pi \mu R V \tag{1}$$

where R is the radius of the particle. V its velocity, and μ the coefficient of viscosity of the fluid. When the velocity of a falling particle is constant, this force must be exactly balanced by the gravitational force acting on the particle, whence

$$6\pi\mu R V = \frac{4}{3} \pi R^3 (\rho_1 - \rho_2)g,$$

$$V = \frac{2 (\rho_1 - \rho_2)gR^2}{\mu} \tag{2}$$

where ρ_1 and ρ_2 are the specific gravities of particles and fluid, respectively, and g the acceleration due to gravity. Stokes' Law has important experimental applications, e.g., in Millikan's determination of the electronic charge by the oil-drop method.

Examination of (2) shows that the velocity of the particle is proportional to the square of its radius. Hence, as the particle size decreases, the velocity of fall decreases very rapidly and the rate of settling of small particles will be almost infini-

* With large particles eddy-currents may be set up and the particle falls in turbulent motion. "

tesimal, though never zero.   In other words, an isolated particle, with no others sufficiently near to influence it, will always eventually reach the bottom of the vessel.   However, as the particle size is decreased, another factor becomes important, the effect of the impacts of the fluid molecules upon the particle.   Since the molecules of a liquid are in a state of constant thermal motion of either a translatory or vibrational character, they continuously bombard the particle.   As long as it is large compared with the fluid molecules, the impacts received from any one direction will be substantially balanced by those received from another, so that their total effect will cancel out and the particle will fall vertically.   Furthermore, the inertia of a large particle will be relatively great and its displacement, due to any given impact by a liquid molecule, negligible.   As the particle becomes smaller it will tend, at any moment, to receive more impacts from one direction than from another, causing it to move correspondingly.   Also, owing to the particle's decreased mass, the effects of these molecular impacts will be much more pronounced and the particle will exhibit an oscillatory or jiggling motion in addition to its vertical fall.   Thus, instead of falling in a vertical, straight line, the particle descends by a broken zig-zag path, but, as long as only a single particle is considered, it will ultimately fall just as surely as though the effects of molecular impacts were disregarded.

The Brownian Movement.

This oscillatory or zig-zag motion is exhibited by all particles whose size falls within the colloidal range.   First observed in 1826 by the English botanist Brown in suspensions of certain spore dusts in water, he showed that it was also exhibited by mineral bodies, *e.g.*, particles of coal dust, and that it was by no means limited to living matter.   Later investigations have shown beyond dispute that the Brownian movement is actually due to unbalanced molecular impacts on the small particle and not to other causes such as convection currents set up in the liquid, electrical effects, vibration, etc.   The phenomena of the Brownian movement afforded the first visual confirmation of the correctness of the kinetic theory of matter.

made by Perrin * and confirmed by Westgren and others, are so convincing that there is little cause to doubt the correctness of the molecular theory.

The Perrin Sedimentation Equilibrium under Gravity.†

If, instead of the single particle, a large number of particles are suspended in a relatively small volume of fluid an additional effect superimposed upon the normal settling will be obtained, due to the interaction of the settling particles upon each other. If, at the beginning of sedimentation, the concentration of particles be uniform throughout the suspension, then, providing thermal convection and other disturbing influences be carefully excluded, a large number of particles will eventually concentrate in the lower layers of liquid, while the upper will be relatively free from suspended particles. The particles concentrating in the bottom will tend to act upon each other in much the same way liquid molecules act in causing Brownian motion. In other words, a particle coming into the region of high concentration will receive many more impacts on its under than on its upper side, due to the movement of the larger number of particles underneath. This will tend to counteract the gravitational forces causing settling, and thus establish a vertical concentration gradient.

Visualization of the formation in a suspension of this concentration gradient caused by Brownian movement may be confused because of the interference of the solvent molecules. The phenomenon may be compared with the behavior of a mixture of two gases, the whole at constant temperature and pressure, but with a concentration gradient of the two components. Consider a plane through such a gas mixture perpendicular to the direction of the concentration gradient. Assume that the concentration of component A is high to the left of this plane and low to the right; the distribution of component B must be the reverse of this. In unit time more molecules of A will impinge on the left side of the plane in question than on the right, the converse being true for molecules of B. Consequently, more molecules of A will pass through the plane towards the right than will move in the opposite direction and, similarly, more molecules of B will travel in the reverse direction. This must result in an equalization of the concentrations of the two components, the process being the phenomenon of gas diffusion.

* PERRIN, J., *Bull. Soc. Fr. Phys.* **3**, 170 (1909).
† PERRIN, J., *Atoms*, tr. D. L. Hammick, Constable and Co., London (1923).

If, now, a liquid suspension be visualized, in which a similar concentration gradient of suspended particles exists, it is clear that a similar argument can be applied to motion of the solid particles and solvent molecules across a plane at right angles to the concentration gradient.   However, the total number of particles per unit volume no longer remains constant and the argument must be modified accordingly.   Clearly, the number of solvent molecules hitting the plane from the direction of higher particle concentration will be less than that from the opposite direction because of the presence of particles barring their path.   Hence, a net transfer of solvent past the plane occurs in the direction of lowered solvent concentration and this can only occur through movement of the solid particles in the opposite direction to make way for the solvent.   The movement is readily explained by the fact that more particles must hit the plane from the direction of high particle concentration.   This diffusional force is balanced by the gravitational pull on the particles downwards, if they are heavier than the liquid.

Smoluchowski * pointed out that the concentration gradient existing in a suspension was exactly similar to the attenuation of the atmosphere with increasing height from the earth's surface and could be calculated in a similar manner.   Consider a vertical column of gas of indefinite height and unit horizontal cross section.   The pressure on any horizontal section is the weight of gas above it, and the rate of change in pressure dP with height dH is therefore equal to the density ρ, i.e., $dP = -\rho g\,dH$. Assuming the gas laws apply,

$$\rho = \frac{1}{V} = \frac{MP}{RT}\,.$$

Substituting and integrating, one obtains the familiar formula for the static isothermal change of barometric pressure with altitude,

$$\ln \frac{P_0}{P} = \frac{MgH}{RT} \tag{3}$$

Inspection shows that the decrease in pressure with height is greater the larger the molecular weight of the gas.   Since, in a gaseous mixture, each component will exert its pressure independently of the other, there will tend to be a variation in composition with the height.†   In solution the solute at suffi-

* SMOLUCHOWSKI, M. V., *Ann. Physik* (4) 21, 756 (1906).

† In the atmosphere this effect is obscured by the mixing action of air currents. However, at the highest altitudes the per cent of oxygen is in fact lower than at sea level.

ciently high dilution exerts an osmotic pressure corresponding to the gas law. One would, therefore, anticipate a corresponding distribution of solute concentration with depth. This is indeed encountered, as indicated above, but the distribution law is in this case modified by the buoyant effect of the solvent.

Consider a cylinder of unit cross-sectional area, height dH, closed at each end with a semipermeable membrane, filled with the dispersion under consideration, and immersed in a bath of the pure solvent or dispersing medium (Fig. 1). Let ρ, ρ_0, ρ_s be the densities of the pure solute, pure solvent, and solution, respectively. Hydrostatic head inside the cylinder $= g\rho_s dH = -dP_s$. Hydrostatic head outside the cylinder $= g\rho_0 dH = -dP_0$, where P_s and P_0 are the hydrostatic pressures of solution and solvent, respectively. At equilibrium there can be no tendency for solvent to enter or leave the cylinder, i.e., no circulation. Hence, the difference in hydrostatic pressures must be balanced by the osmotic pressure of the solute particles π.*
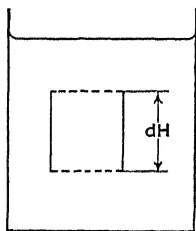
FIG. 1.

$$dπ = dP_s - dP_0 = g(\rho_0 - \rho_s)\,dH. \qquad (4)$$

If w is the concentration of the solute in the dispersion, in mass per unit volume, then, assuming volumes additive on mixing,

$$\rho_s = \rho_0 + w\left(\frac{\rho - \rho_0}{\rho}\right).$$

$c = w/M$, where c is the molal concentration and M the molecular weight of the solute particles.† Also, $\pi = cRT$ (assuming

---

* It must be kept in mind that π is the pressure necessary to prevent the solvent's going through the membrane.

† This applies even to coarse suspensions, in which the solute particles are, in many cases, undoubtedly mere physical aggregates. In the ultimate analysis, the only method available to the chemist for determining the size of excessively fine particles is by determination of their pressure-temperature-volume relation. This may be done by measuring gas density or osmotic pressure or its equivalent in solution, followed by interpretation by the use of Avogadro's hypothesis. Of itself this technique throws no light whatever on the nature of the forces holding the ultimate particle together. In this sense, any suspension, however coarse, provided it is sufficiently stable to enable the determination at a given concentration of its osmotic pressure or equivalent, may be said to have a molecular weight which corresponds to its particle size. In many cases there are sound chemical grounds for believing that the forces holding the particle together are exclusively of the valence type; in other instances, there is equally convincing evidence that the

the osmotic pressure follows the gas laws), whence $d\pi = RT\,dc$. Algebraic combination gives

$$\frac{dw}{w} = \frac{dc}{c} = (\rho_0 - \rho)\frac{Mg}{DRT}\,dH.$$

Integrating

$$\ln\frac{w_1}{w_2} = \ln\frac{c_1}{c_2} = \left(1 - \frac{\rho_0}{\rho}\right)\frac{MgH}{RT} \tag{5}$$

where c_1, c_2 are the concentrations in mols at zero height and H, respectively.* Also, $c = n/N$, where n is the number of particles in 22.4 liters and N is the Avogadro constant. Hence,

$$\ln\frac{n_1}{n_2} = \left(1 - \frac{\rho_0}{\rho}\right)\frac{MgH}{RT}. \tag{6}$$

Provided the disgregating tendency of a suspension of non-agglomerating particles, of whatever size, is quantitatively equal

---

particles are held together, in considerable part at least, by purely physical attractive forces (which may, of course, include the intermolecular attractions). In many cases it is difficult to distinguish between the two types of attractions. It is, however, important always to keep clearly in mind the fact that the particle size determination itself throws no light on the nature of the intra-particle forces.

* If volumes are not additive on mixing, one must write

$$\rho_s = \rho_0' + w\frac{\rho' - \rho_0'}{\rho'},$$

the differential equation becoming

$$\frac{dw}{w} = \frac{gM}{RT\,\rho'}(\rho_0' - \rho')\,dH$$

where ρ_0' and ρ' are the partial densities of solvent and solute, respectively, in the dispersion. Integration requires knowledge of ρ_1' and ρ' as functions of w or c. If over a given range variation in the partial densities may be neglected, the expression integrates into a form which simplifies to the above equation provided $\rho_0' = \rho_0$ and $\rho' = \rho$.

Equation (3) for gases can also be derived from (4). If P is the partial pressure of the gas to which the piston is impermeable and P_0, P_0' are the partial pressures of the other gas outside and inside the cylinder, respectively, equating the differences in total densities inside and out,

$$\frac{(\rho_0 - \rho_s)}{g} = \frac{M_0\,P_0}{RT} - \frac{M_0\,P_0'}{RT} - \frac{MP}{RT}.$$

Granting $P_0 = P_0'$, as is necessary to prevent continuous circulation,

$$\rho_0 - \rho_s = -g\frac{MP}{RT}$$

where P corresponds to the osmotic pressure, π, in the case of liquids. Hence

$$dP = -g\frac{MP}{RT}\,dH$$

$$\ln\frac{P_0}{P} = \frac{MgH}{RT}.$$

to the pressure calculated from the gas laws, these equations make it possible to compute concentration distribution under gravity (or any comparable field of force). The simultaneous presence of solvent molecules does not interfere because their effects cancel out quantitatively. The indications are that, even at high concentrations, there is no net interference of solvent particles with the thermal disgregation tendency of a material of whatever particle size.*

If one assume a progressive increase in the molecular weight of the solute, *i.e.*, a progressive growth in the size of its ultimate particles, the equations indicate a rapidly increasing concentration gradient of solute with depth of solution, but one would none the less expect the particles to remain in suspension, if the disgregating force holding them apart, the thermal agitation of the Brownian movement,† be adequate. It is found experimentally that one can obtain suspensions of solid particles in liquids, which behave exactly as here indicated, even though the size of the ultimate particles is enormous in comparison with that of ordinary molecules with which the chemist is commonly dealing. Thus, Perrin verified (6) by the use of gamboge ‡ suspensions in water. Such suspensions are readily obtained by rubbing the solid under water with a soft brush. By fractional centrifuging, suspensions were prepared in which the particles were of substantially uniform size. A column of such a suspension, 0.1 mm. in height, was observed by a microscope equipped with a dark-field condenser (p. 134). Focusing at different levels, the number of particles at each level were counted after equilibrium had been reached, no change occurring after three hours, even up to fifteen days. Typical results are given in Table II, showing substantial agreement between observed and calculated values.

Westgren,§ using gold sols over a much wider range of experimental conditions, obtained even closer agreement with (6) for a total height of over 1 mm. It is interesting to note that he

---

* One must not make the error of applying this conclusion to rate of disgregation. The interference of solvent molecules with rate of movement of the dispersed particles is often very great indeed, *e.g.*, if the solvent has a high viscosity.

† If the particles become very large, molecular impacts will never supply sufficient energy to exceed the minimum required for the motion of the particles, for example, rocks at the bottom of the sea.

‡ Gamboge is a gum made from the dried milk of a guttiferous plant.

§ WESTGREN, A., *Z. physik. Chem.* **89,** 63 (1915).

reached the same equilibrium either by allowing the solute to settle or to diffuse up from a concentrated layer at the bottom of the observation slide.

TABLE II.   NUMBER OF PARTICLES AT DIFFERENT LEVELS
IN A GAMBOGE SUSPENSION*

h (in μ)	Observed n	Calculated (Eq. 6) n
0	100	——
25	116	119
50	146	142
75	170	169
100	200	201

If one can measure directly the mass of the individual particle by any suitable technique, since the value of M is clearly Avogadro's constant N times this mass, one has at hand a method for experimental determination of N.   This method was used by Perrin, e.g., who measured the diameter of the individual spherical particles by visual observation in the microscope after precipitation, calculated the volume V, of the single particle, and its mass from its known density ρ.   Substituting in (6) for M its value $NV\rho$, one obtains

$$\ln \frac{n_1}{n_2} = \frac{NV}{RT} (\rho - \rho_0)gH. \qquad (7)$$

In a large number of experiments Perrin found $N = 6.5$ to 7.2×10^{23}.   Westgren † obtained even closer agreement with Millikan's ‡ generally accepted value, 6.06×10^{23}, namely, 6.05×10^{23}.

From the foregoing, it would be expected that the concentration of an ordinary colloidal dispersion would become enormous at a depth of even one centimeter.   In practice no such sedimentation is observed, the concentration, e.g., of a gold sol in a beaker, being substantially uniform throughout.   This non-sedimentation of normal colloidal dispersions is probably due to mixing effects of convection currents.   The velocity of fall of

* PERRIN, J., Compt. rend. 146, 967 (1908).
† WESTGREN, A., Z. anorg. Chemie 93, 231 (1915).
‡ MILLIKAN, R. A., Phys. Rev. 35, 1231 (1930).

particles in the colloidal range is almost vanishingly small, and hence the slightest eddies would suffice to prevent sedimentation.

Sedimentation in a Centrifugal Field.

By increasing the force under which settling takes place, as by centrifugal action, rate of settling can be increased or, alternatively, finer particles can be caused to settle, gravity being insufficient for particles less than about 50 or 100 mμ. A relation similar to (2) is readily deduced, using Stokes' Law. If a particle of radius R settles under the action of centrifugal force through a radial distance dx, in time dt, under an angular velocity ω, and volumes can be assumed additive, then, equating Stokes' Law of friction to the net centrifugal force,

$$6\pi\mu R\frac{dx}{dt} = \frac{4}{3}\pi R^3(\rho - \rho_0)\omega^2 x,$$

or

$$\frac{dx}{x} = \frac{2}{9}R^2(\rho - \rho_0)\left(\frac{\omega^2}{\mu}\right)dt.$$

Integrating between the distances, x_1 and x_2,

$$\ln\frac{x_2}{x_1} = \frac{2R^2(\rho - \rho_0)\omega^2 t}{9\mu}. \tag{8}$$

This is the basic principle of various centrifuges on the market.

Svedberg has taken advantage of centrifugal force in his sedimentation methods for determination of the molecular weight of colloidal materials. The ultracentrifuge consists essentially of a rotor M (Fig. 2), driven by twin oil turbines T, carrying the transparent cell C in which is placed the dispersion to be studied. The degree of sedimentation can be measured either by colorimetric or refractometric means, or by absorption of ultraviolet light. In the latter case, a beam of light L is passed through the cell to a camera P, exposures being made by the electromagnetically controlled shutters E_1, E_2, which admit the light only during the short time the cell is in the beam in each revolution. The change in opacity with depth can be evaluated photometrically by comparison with standards. Centrifugal fields equivalent to 900,000 times gravity have been attained,

routine measurements being possible in fields of 700,000 times gravity. It is clear that to avoid vibrational or thermal effects at speeds necessary to produce such fields, elaborate precautions are necessary. The machine must be designed with great care,
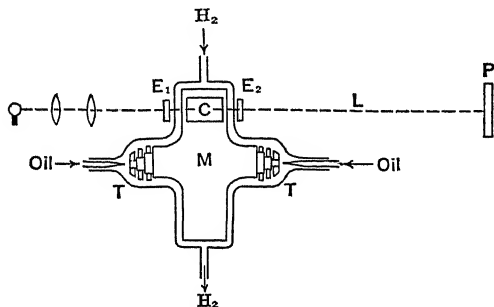
FIG. 2.   The Ultracentrifuge.

the bearings, for example, offering a particular problem.* To avoid thermal convection currents, hydrogen at 15 to 25 mm. pressure is circulated through the rotor housing. For lower fields, 500 to 15,000 times gravity, the rotor is driven directly by an electric motor and hydrogen at atmospheric pressure is employed. Of late cheaper air-driven centrifuges have been designed but lack the easy control and great resolving power of the Svedberg instrument.

Two types of measurements are made with the ultracentrifuge. The first, sedimentation equilibrium, is analogous to the Perrin sedimentation under gravity, but may require days for equilibrium to be set up between the diffusional and sedimentational forces. Visualize the tube of a centrifuge of unit cross-sectional area,† rotating at an angular velocity of ω radians per second, and containing a solution the concentration of which, at the distance x from the center of rotation, is c gms. of solute per c.cm. The density of the pure solvent is ρ_0 and that of the pure solute ρ gms. per c.cm. For the moment, assume the volumes additive on mixing. Focusing attention on the differential cell of liquid dx, the radius of rotation of which is x, the

* SVEDBERG, T., and PEDERSEN, K. O., *The Ultracentrifuge*, Clarendon Press, Oxford (1940).

† A differential area may be assumed, but will subsequently cancel out.

mass of solute in the cell is *cdx* and hence the centrifugal force operating on the solute is $c \cdot dx \cdot \omega^2 \cdot x$.   However, a corresponding centrifugal force operates on the solvent, causing the latter to buoy up each solute particle, so that the net centrifugal effect on the solute is obtained by multiplying the total centrifugal force acting upon it by the term, $(\rho - \rho_0/\rho)$.   Since the cell is of unit area the net centrifugal force operating on the solute, $c \cdot dx \cdot \omega^2 \cdot x \cdot (\rho - \rho_0/\rho)$, represents a pressure tending to drive the solute radially outward.   This pressure must operate against the disgregating tendency caused by the osmotic effect.   Because, assuming the gas laws, the osmotic pressure is equal to the product of RT and the molal concentration, the osmotic pressure effective in counterbalancing the pressure due to centrifugal force is $d\pi = (RT/M)\, dc$.   When sedimentation equilibrium is reached, these two pressures can be equated, giving

$$M \cdot \omega^2 \cdot x \cdot \left(\frac{\rho_0 - \rho}{\rho} \right) \cdot dx = RT\, d \ln c.$$

Integrating between limits and rearranging,

$$M = \frac{2RT \ln \frac{c_2}{c_1}}{\omega^2 \left(\frac{\rho - \rho_0}{\rho} \right)(x_2{}^2 - x_1{}^2)}. \qquad (9)*$$

Study of the derivation makes it clear that this relation is independent of such factors as nature of the solvent, shape of the particle, and the like but, in common with the classical methods of determining molecular weight in solution, it assumes applicability of the gas laws to osmotic pressure.   Indeed, it is equivalent to a great expansion of the range of the classical methods, which lose precision at molecular weights of 3000 to 5000. Molecular weights in the millions have been determined by the ultracentrifuge.

In the second method of using the centrifuge, the rate of travel of the moving boundary of a dispersion into clear solvent is measured.   The method is rapid but requires independent

---

* It is clear that, where volumes are not additive on mixing, partial densities should be employed, but this correction is often of minor importance.

determination of the diffusivity of the solute in the solvent in question.   It is the one most frequently used.

The Translatory and Rotary Brownian Movement.

Molecular bombardment of the particles in a dispersion should result in their completely random motion and experimentally this is found to be the case.   The zig-zag, random travel of a particle is known as the translatory Brownian movement.   A typical tracing of the movements of gamboge particles over a period of time is shown in Fig. 3.   In order to form a basis for calculation, a distance Δx is defined, which is the square root of the average of the squares of the individual displacements Δx_1, Δx_2, etc., of the particle along the x axis in equal times t, $i.e.$, the so-called root mean square of the displacements.   Einstein* and Smoluchowski,† working independently, obtained for this average distance Δx, moved through by a particle of radius r in a liquid of viscosity μ during a time t, the theoretical expression
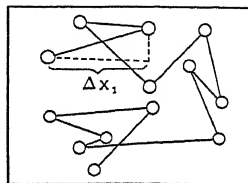
Fig. 3.

$$\Delta x^2 = \frac{RT}{N} \frac{t}{3\pi \mu r} \tag{10}$$

where T is the absolute temperature.   Svedberg, Perrin, and others succeeded in confirming this relation experimentally, Perrin obtaining an average value for N of 6.8×10^{23}.

Besides the translatory motion, the particles also possess a rotary motion, given by the formula, due to Einstein,‡

$$\alpha^2 = \frac{RT}{N} \frac{t}{4\pi \mu r^3} \tag{11}$$

where α is the root mean square of the angle of rotation in time t.   Perrin § verified this equation by watching the motion of comparatively large mastic grains of 6.5 μ radius, determining rate of rotation by following under the microscope defects on

* Einstein, A., *Ann. Physik* (4), **17**, 549 (1905).
† Smoluchowski, M. v., *Loc. cit.*
‡ Einstein, A., *Ann. Physik* (4), **19**, 371 (1906).
§ Perrin, J., *Compt. rend.* **149**, 549 (1909).

the surface of particular globules, obtaining a value for N of 6.5×10^{23}.

The Stability of Dispersed Particles against Agglomeration.

From the above it can be seen that, if the particles are sufficiently small, Brownian movement may keep them in suspension against gravitational forces, providing they do not cohere on impact. It remains to explain why the larger particles do not grow at the expense of the smaller ones, and why the particles, which are all in rapid thermal motion, do not coalesce on collision.

Just as the vapor pressure of small drops is greater than that of the liquid in bulk, the solubility of small spherical particles will be greater than that of large. For this case, equation (8) of Chapter V becomes

$$\frac{RT\rho}{M} \ln \frac{S_r}{S} \qquad \frac{2\sigma}{r} \qquad (12)$$

where S is the solubility in bulk, *i.e.*, at a plane surface, S_r that at a surface of curvature r, and σ the interfacial tension. The equation is readily understood, since in subdividing the material into spheres of radius r, a large amount of new surface is brought into existence and hence work must be done against the interfacial tension. Equation (12) affords one of the few methods by which the surface tension of a solid may be estimated.* Theoretically, it is applicable only to spheres, and hence in effect only to liquids or to amorphous solids. None the less, there are well-recognized parallelisms of behavior of crystalline particles, *e.g.*, increasing solubility with decreasing particle size, for which this equation probably offers the best method of approximation. In all dispersions of solids, the smallest particles should dissolve continuously in the solvent and redeposit on the larger ones, until finally a coarse dispersion, capable of settling out, is obtained.

The experimentally observed behavior of solid suspensions indicates that, even when dealing with materials so insoluble that solution cannot be looked upon as an effective mechanism

---

* For example, DUNDON, M. L., *J. Am. Chem. Soc.* 45, 2479, 2658 (1925).

of change, the particles frequently exhibit a strong tendency to agglomerate in groups or clumps which, once formed, are frequently quite stable. Granting that the surfaces of two solid particles come into contact, the surface fields of force would, in general, be of sufficient intensity to require a considerable expenditure of energy to effect separation. In a suspension the only energy available for separation is that of Brownian movement and it is not surprising to find that this is often insufficient to separate even particles of extremely small size.*

One frequently encounters suspensions of inert, solid particles in relatively pure liquids, which show little or no tendency toward agglomeration. Investigation of such systems indicates that the suspended particles usually carry an electric charge relative to the liquid suspending medium. Thus, if the suspension be placed between direct current electrodes, the particles are observed to travel with the current to one or the other electrode. Furthermore, the mobility of the particles in an electric field is approximately the same as that of ordinary ions; e.g., the mobility of silver particles in a silver dispersion is found to be approximately 3.7×10^{-6} cm. per sec. per volt per cm., whereas that of the sodium ion is 4.5×10^{-6}. McBain and Williams † carried out experiments in which oil globules dispersed in water were subjected to an electric field acting in the reverse direction to the gravitational pull on the particles. When the drops failed to rise or fall, the electrical pull Fe_n was equated to the gravitational pull (in appropriate units)

$$Fe_n = \frac{4}{3} \pi r^3 (\rho - \rho_0). \tag{13}$$

It was found, for example, that an oil globule of radius 0.0041 cm. carried 5.9×10^5 electronic charges, although the proportion of the surface covered was only about 0.0006 of the total area. Although these results were obtained for oil drops, without doubt comparable results would be obtained with solid particles, the number of unit electronic charges per particle being

* However, if particle size gets sufficiently small, separation should become possible. There can be no doubt that there is a general tendency for dispersion to result as particle size decreases, i.e., dispersability tends to go up as molecular weight goes down.

† McBain, J. W., and Williams, R. C., Colloid Symposium Annual 7, 105 (1929).

large but the ratio of charge to surface area or mass small.
From this point of view the colloidal particle may be pictured
as an extremely large polyvalent ion.*   Since the colloidal solu-
tion is electrically neutral, there must be in the liquid external
to the particles a total charge equal in magnitude and opposite
in sign to that on the particles.   This charge must be in the
form of ions and is the important factor in stabilization of the
particles (see pp. 125-131).

In sharp contrast to dispersions in which the particles are
electrically charged is a class of colloidal solutions, the proper-
ties of which will be taken up in detail in the next chapter.   In
these, stabilization appears to be effected by solvation of the par-
ticles.   This class of colloids is known as emulsoidal, lyophilic,
or hydrophilic,† while the class stabilized by an electric charge
is known as suspensoid, lyophobic, or hydrophobic colloids.

The Preparation of Suspensoids.

It has been emphasized that one can obtain dispersions of
particles the size of which varies from atomic ‡ to visible dimen-
sions.   Where the particles are in the lower end of this size
range, the dispersions are called solutions and their properties
are treated by the ordinary techniques of physical chemistry.
Where the particles are very large, the suspensions are entirely
unstable, settling out so rapidly that their study as a dispersion
is impracticable.   Between these limits of particle size are found
large numbers of suspensions, many extraordinarily stable, and
all possessing properties differentiating them sharply from sys-
tems outside the range.   These intermediate suspensions, called
sols, are described as colloidal (see p. 96).   It is obvious that
there can be no sharp transition between true solutions and
colloidal suspensions at the lower limit, or between the latter

---

* The precise mechanism of the genesis of the charge is at this stage immaterial
(see p. 119).

† These words are derived from Greek roots, "hydrophilic" and "hydrophobic"
meaning "water loving" and "water hating," respectively. "Lyophobic" and "lyo-
philic" correspond to Greek roots "fear to loosen" and "love to loosen," but by analogy
with hydrophilic and hydrophobic have come to mean "solvent hating" and "solvent
loving" in modern colloid terminology. They are therefore equivalent to suspensoid and
emulsoid, respectively.

‡ It is not intended to imply that fine dispersions are generally atomic. However,
the word *molecule*, as frequently used, covers such an enormous range of dimensions that
to use it to indicate fineness of particle size is ambiguous.

and gross suspensions at the upper. None the less it is found in practice that the transition from solutions to sols occurs at particle dimensions of approximately 1 mμ, and from sols to suspensions at about 1 μ. It is to be expected that possession of colloidal properties will depend more upon particle size than upon the characteristics of the dispersed material. While the facts indicate that any substance can exist in the colloidal condition, it must be kept in mind that its properties will profoundly influence the stability of the colloidal state; *e.g.*, high volatility or solubility induces particle growth and makes a stable colloidal condition impossible.

Two general methods of preparing suspensoids may be distinguished, disgregation and condensation (in which particles of molecular dimensions are built up until they reach colloidal dimensions). Of these, the latter has generally been the more successful.*

Mechanical grinding is the most obvious method of dispersion. Many substances disperse on wet or dry grinding, but an enormous energy expenditure is required to grind them down completely to colloidal dimensions.† Various types of mechanical mills,‡ so-called "colloid mills," have been devised for the purpose. Most of them have smooth grinding surfaces and work upon the principle illustrated in Fig. 4, the coarse particles suspended in the dispersion medium being disintegrated primarily by high shearing forces.

FIG. 4.   Section through a Colloid Mill.

Despite their inefficiency in subdividing compact solid particles, they are frequently effective in breaking up or redispers-

---

* In many methods involving dispersion, closer inspection shows the condensation phase of the operation to be the more important.

† This can be seen from Kick's Law,

$$E : b \ln \frac{L_1}{L_2}$$

where E is the energy required to crush a material from an initial linear dimension L_1 of the particles to a final linear dimension L_2, b being a constant.

‡ For use in preparing emulsions, see p. 254.

ing aggregates or flocculates of small primary particles, even where these may be relatively firmly held together.

The preferable technique of preparation of a colloidal suspension consists in building up particle size by some form of precipitation or agglomeration of the highly dispersed material, under conditions making it practicable to control the size of the solid particles synthesized. Thus, colloidal gold is prepared by vaporization of the metal in an electric arc struck between gold electrodes under water. The vapor condenses and collapses to particles, which are necessarily small because the amount of gold vapor in the immediate neighborhood of any given point of condensation is limited (see p. 118). Alternatively, the sol can be prepared by reduction of a solution of a gold compound, such as the chloride. This again must be done under conditions limiting the amount of gold which can accumulate in a given particle and the possibility of continued growth of the particle once formed. A third effective technique involves precipitation of a dilute solution of a solid material by addition of a liquid miscible with the solvent, but in which the material itself is insoluble, as illustrated in the preparation of Perrin's gamboge suspension. Despite the apparent diversity of such techniques, they are in all cases governed by important common principles.

From inspection of (12), keeping in mind the fact that it is applicable, at least qualitatively, to crystalline particles of a given material, it is clear that in order to initiate any separation of discrete particles by precipitation from a solution, the solution must be supersaturated with respect to the bulk solubility of the material involved. This is because, even though a small particle should succeed in forming at a given point, a solution of concentration S is still unsaturated with regard to the particle, so that it would immediately start to redissolve, and consequently disappear. Furthermore, the factor determining the stability of a particle of given size, from this point of view of its capacity for continued existence, is the relative saturation S_r/S, or, equivalent to the same thing, the relative supersaturation, $(S_r - S)/S$.

Consider a solution of the material of very low solubility S, the supersaturation being enough to induce formation of a few

particles sufficiently large to resist re-solution. Initial formation of each such particle will deplete the solution immediately surrounding it of solute but, since it is assumed stable despite this fact, as solute diffuses in toward its surface from the more concentrated solution at a distance, the increase in concentration of the liquid in actual contact with it will induce continued growth.* If left undisturbed, this growth will continue until the crystallite comes into complete equilibrium with the solution. Since only a few crystallites are formed in a given volume of liquid and these pick up all the solute separated from the solution by particle growth, the particles formed will be relatively large and the solution concentration will fall to substantially bulk saturation. The particles initially formed, which can serve as centers for further crystal growth, are known as nuclei. They may be so small that they are undetectable by ordinary means, sometimes revealing their presence only by growth under properly chosen conditions.

If the solution initially taken has a higher concentration, the number of nuclei forming per unit volume and time, expressed as a statistical average, will certainly increase. If this increase in nucleus formation per unit increase in supersaturation be sufficiently high, the concentration of the solution can be so depleted by the consumption of material in the synthesis of the nuclei themselves that further growth by direct precipitation is impossible. This means that the nuclei or crystallites are now stable from the point of view of size, except as the larger particles may grow at the expense of the smaller ones by the process of re-solution of the latter followed by precipitation on the former. In other words, solutions of high supersaturation can precipitate in a relatively stable colloidal size range.† Since both rate of solution and rate of crystallization are proportional to concentration differences, whereas differences in particle size determine relative concentrations, the higher the bulk solubility of the material, the less stable it is in colloidal form, because of more ready growth of large particles at the expense of small ones.

* The linear rate of growth of a particle is proportional to absolute supersaturation, rather than relative. It may differ on the various faces of a crystal.

† Particle growth can be induced by maintaining a low degree of supersaturation, as in sugar crystallization.

Even in the case of a substance with marked tendency toward crystallization, particles sufficiently small in size are either too inadequately developed as crystallites to have their crystalline properties readily recognized, or perhaps are too small for complete development of crystal structure.   In any event, they frequently exhibit amorphous characteristics in high degree. Furthermore, if the solution be very concentrated, precipitation can induce the development of an enormous number of nuclei, and consequently of amorphous particles, per unit volume.   As will appear later (see p. 221), high concentrations of finely divided particles suspended in a liquid are very prone to develop well-defined so-called gel structures, *i.e.*, bodies exhibiting quasi-fluid properties, particularly abnormal resistance to flow at low shearing stresses (see, however, Chapter VIII), so that gelation can be induced almost at will by this technique of control.

These phenomena were first studied comprehensively and the relationships clearly enunciated by von Weimarn.*   He assumes the rate of nucleus formation per unit volume to be directly proportional to relative supersaturation for any given solute, solvent, and temperature.   Only high relative super-saturations can give a colloidal suspension or gel.   Only concentrated solutions can give a gel, since only these have enough solute per unit volume to synthesize the large number of colloidal particles required.   Only materials of low solubility can form stable gels or sols.   Thus, precipitating barium sulfocyanide with manganous sulfate, he found no precipitation whatever, even in a year, with reagent concentrations of 0.00005 to 0.00014 N, corresponding to relative supersaturations up to 3. A tenfold increase of reagent concentrations, giving relative supersaturations up to 48, gave slow precipitation, with development of large crystals, the growth of which, however, required in some cases a matter of months.   Reactants up to 0.75 N, with relative supersaturations up to 21,900, gave rapid precipitation, with crystalline character of the particles becoming less marked as concentrations increased.   Concentrations up to 3 N, a supersaturation of 87,500-fold, gave immediate formation of an "amorphous precipitate," while for 3 to 7 N solutions (super-

saturations up to 204,500-fold) clear cellular jellies were found. This behavior is typical, although complicating factors are frequently encountered.

Condensation may be achieved by interaction of any two substances in a mutual solvent, in which one of the reaction products is insoluble; *e.g.*, by reduction, oxidation, dissociation, hydrolysis, or double decomposition. A well-known example, one of the earliest studied, is the reduction of gold trichloride solutions by various reagents. Faraday used ethereal solutions of phosphorus as the reducing agent. At present, formaldehyde, hydrogen peroxide, and hydrazine are widely employed. The size of the gold particles is regulated by the proportions of the reagents or by precipitation as particles of a previously formed sol under conditions precluding new nucleus formation. Gold sols with the smallest diameter particles are red, while those in which the particles are large are blue in color, due to different light absorption.* Zsigmondy † is responsible for the so-called nucleus method of producing gold sols of nearly uniform particles of any desired size. If to a dilute reduction mixture a certain amount of already reduced gold sol containing nuclei is added, new nuclei formation can be avoided and the gold is precipitated only on those nuclei already present, so that the size of the particles formed can be varied by controlling the number of nuclei added. Among other metals which are readily produced in sol form by the reduction of a suitable salt are copper, silver, platinum, mercury, and tellurium.

Interesting sols are prepared by oxidation. Most important is probably the sulfur sol formed by the reaction,

$$2 H_2S + SO_2 \rightarrow 3 S + 2 H_2O.$$
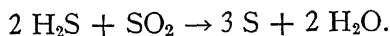
As would be expected, Odén ‡ found, at high concentrations, a higher percentage of colloidal sulfur and more particles of small

---

* When the particles of a sol are of the order of magnitude of the wave-length of light they will absorb certain portions of the spectrum quite completely, giving a pronounced color to the sol. For extremely fine particles, the absorption is mainly in the violet end of the spectrum so that the color of the sol is a very dark red. As the particle size is increased, the absorption shifts to the red, so that the sol progressively changes color, shifting toward the violet and for large particles to black. The color of the sol, then, is an indication of the particle size and, in the case of gold sols at least, is often used to characterize it.

† ZSIGMONDY, R., and SPEAR, E. B., *Chemistry of Colloids*, Wiley (1917), p. 91.

‡ ODÉN, S., *Nova Acta (Upsala)* (4), 3, No. 4 (1913).

diameter.   A well-known example of a sol prepared by hydrolysis is ferric oxide, produced by pouring a solution of ferric chloride into a large volume of boiling water,

$$FeCl_3 + 3 H_2O \rightarrow Fe (OH)_3 + 3 HCl.$$
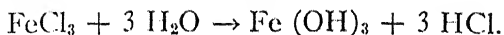
As already stated, another important condensation method is the addition of a solution of the substance to a second liquid, miscible with the solvent, but in which the substance is insoluble.

The condensation technique, due to Bredig,* which consists in striking an arc between electrodes composed of a metal, such as gold, platinum, silver, or the like, under water or an organic liquid (see p. 114), throws light on the stabilization of suspensoids.   Disintegration of the metal of the electrodes is achieved by vaporization, but the colloidal particles themselves are formed by the condensation of the vapor.†   Either direct or oscillating arcs can be used, very high frequency (10^5 to 10^7 cycles) giving more finely dispersed sols.   Electrode distance is often controlled automatically, and the sol is stirred and cooled. Considerable decomposition of the solvent takes place, to overcome which Svedberg encloses the arc in a quartz tube, the metallic vapor being blown out into the solvent through a small hole opposite the arc by a stream of nitrogen.‡

The Mechanism of the Stabilization of Suspensoids.

It has been repeatedly observed that stable, well-characterized sols are never formed by the Bredig method when precautions are taken to insure high purity of the metal wire and the water in which the arc is struck, and to use a chemically resistant container for the solution.   Thus, if one attempts to make a gold sol in conductivity water, it is found that, while a fugitive blue sol may be formed, addition of an electrolyte before sparking is required to obtain the stable, red gold sol.   A certain amount of electrolytic impurities seems necessary for the stabilization of these sols.   Thus HCl can be used for stabilization,

---

* BREDIG, G., *Z. angew. Chemie* 11, 951 (1898).
    † That this is the mechanism is shown by the fact that sols of the alkali metals can be prepared by vaporizing them, mixing with benzene vapors, and allowing the vapors to condense simultaneously at low temperatures.
    ‡ In its original form, Bredig's method is, of course, limited to metals of high melting point, which are not fused under the conditions of experiment.

the amount required being proportional to the gold concentration for a given particle size.   If the arc be struck in water that contains the necessary amount of electrolyte, it is found that the conductivity of the resulting sol is less than that of the original solution.   Moreover, when a gold sol is placed in a strong electric field, the particles move to the positive pole, showing them to be negatively charged.

These facts provide convincing support for the theory that the suspensoidal particles carry an electric charge, responsible

(a)

(b)

(c)
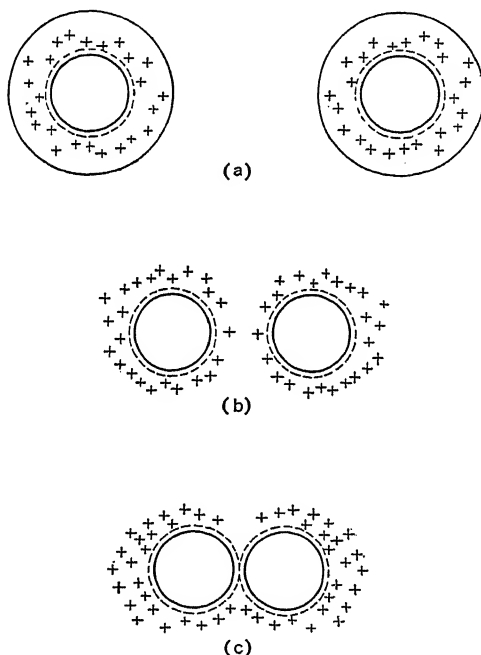
FIG. 5.   Gold Particles.

for their stability.   As to the source of this charge, it is uncertain whether the solid particle adsorbs an ion of a certain charge and thus acquires an ionic atmosphere, or whether it reacts in a chemical sense with the added electrolyte to form a sort of surface complex which can ionize and give the same effect. Whatever the mechanism, the condition of the stable gold particle in a Bredig sol may be conveniently represented as in Fig. 5(a).   Since the gold particle travels to the anode when

placed in an electric field, it may be assumed that it has preferentially adsorbed chlorine ions from the solution, thus localizing a number of negative charges in the immediate neighborhood of the particle's surface.

Granting that the ions on the gold particle are held by surface valence forces, their adsorption is readily accounted for on the Langmuir theory.   Because the particles are formed very rapidly from a highly supersaturated vapor or liquid, the gold atoms are deposited in a haphazard fashion and micro-crystals, if present, will not be arranged in an orderly manner.   Thus the surface atoms will doubtless possess even stronger, unsatisfied secondary chemical valencies than a normal gold surface.   With these secondary valencies the chlorine ions can combine, the mode of combination being immaterial so long as it results in the presence of a negative charge.   The hydrogen ions may be visualized as associated with the adsorbed chlorine ions but, as indicated in Fig. 5(a), dispersed in a diffuse manner around the particle at a certain average distance from it.   They constitute the so-called contra-ions of the system.   The concentration of the hydrogen ions is apparently only slightly greater around the particle than in the main body of the solution.   This segregation of the charges around the particle in two relatively distinct layers resembles to a certain extent the distribution of charges on the plates of a condenser and has been called the electric double layer.   It differs from a condenser, however, in that the charges on one plate are not fixed but are free to move. This picture of the particle in a gold sol, even though imperfect, is useful in studying its behavior.*

Consider a gold sol as pictured in Fig. 5(a), assuming the gold particles negatively charged and the contra-ion hydrogen. Around each gold particle one can draw a surface enclosing it and the contra-ions, the total of whose charges is equal to its own.   The solid thus enclosed is, as a whole, electrically neutral,†

---

* Undoubtedly, in some sols the particles have much more complex structures; thus, the platinum sols, prepared by arcing, are generally considered to consist of a nucleus xPt, yPtO$_2$, stabilized by the adsorbed anion of hexahydroxy-platinic acid H$_2$Pt(OH)$_6$, the hydrogen ions being the contra-ions.   In every case, however, the particle possesses an electric double layer, which governs its characteristic properties.

† There is, of course, a certain field of force around it because of the localization of the charges near its surface.

so that gold particles, separated by a distance greater than the diameter of this surface, will have no mutual electrical effect upon each other. However, visualize two gold particles, which happen to be approaching owing to their Brownian movement, coming close enough together so that their contra-ionic surfaces begin to intersect. The two gold particles now begin to repel each other because of their negative charges, the effective repulsion increasing as they approach, not only because of the decrease in distance, but because, when close together, there are few if any contra-ions between them to interfere with their mutual electric repulsions. The progressive stages of such an approach are indicated schematically in Figs. 5(a), (b), and (c). The mutual repulsion of the two particles, incident to their charges of similar sign, is a force opposing further approach. If this force, operative through the distance involved, is sufficient to absorb their kinetic energy of approach, they will be brought relatively to rest before impact and collision. They can never touch unless approaching with an effective kinetic energy, which is above a minimum value, determined by the magnitude of their charges and the distribution of the contra-ions. Without a certain degree of interparticle contact agglomeration is impossible. Clearly, therefore, an adequate electric charge on the particles of a suspensoid, sufficiently fine to be in active Brownian movement, will maintain it as a stable sol.

Coagulation of Suspensoids.

Visualize a suspensoid with a certain electric charge on the particles and definite conditions of concentration of the contra-ions in the surrounding liquid, such that, in general, contact of the particles will not occur. The translational kinetic energy varies not only from particle to particle, but also, for a given particle, with the time (Fig. 1, Chapter I). The energy distribution corresponds, at least approximately, to the probability law. Occasionally two particles will be approaching each other with velocities and corresponding kinetic energies far in excess of the average values of these quantities characteristic of the particles of the suspension in question. It is easily possible that, whereas the average kinetic energy of the suspensoid

particles is far from sufficient to enable two colliding particles
to overcome their mutual electrical repulsions, those particles,
possessing abnormally high kinetic energies, corresponding to
the extreme right hand side of Fig. 1, Chapter I, may be able to
overcome the electric repulsions and come into contact. Fur-
thermore, at the instant of contact, velocity of approach will
have been greatly reduced, so that tendency to rebound and
separate is a minimum. The surface forces now in operation
tend to hold the particles together so that, in the absence of
other disgregating effects, permanent coagulation occurs. In a
suspension in which the repulsions due to charges do not greatly
exceed the average kinetic effects of Brownian movement, the
probability of rapid approach is great, since a relatively large
number of the particles possess velocities sufficient to place them
above the translational energy requisite for contact. Coagula-
tion, therefore, will be correspondingly rapid. In another sus-
pension, in which the charges are larger, the probability of ap-
proaching particles possessing the necessary energy level is
smaller and the stability correspondingly greater. Even here,
however, occasional contacts on collision can be anticipated.
Such is apparently the condition of the Faraday gold sols
already referred to.

Increasing temperature increases the average kinetic energy
of translation without affecting the percentage distribution of
Fig. 1, Chapter I. Hence, higher temperature increases the
probability of a given particle possessing an energy level suffi-
cient to overcome the effect of the electrical forces. In general,
temperature change has relatively little effect on ionization and
ionic distribution. Consequently, one would expect increase of
temperature usually to increase rate of coagulation. This is,
in fact, found to be the case. However, the rate of rise of
kinetic energy with temperature is relatively low so that the
general influence of temperature on coagulation is small in
comparison with other more important influences.*

Granting the mechanism of particle stabilization by charge
developed above, one would expect ionic conditions in the solu-
tion to be of extreme importance, influencing both charge on the

* It is impossible completely to disentangle all factors affecting coagulation.

particle and the electrical conditions in the surrounding liquid. It is important to have a clear concept of the relevant facts.

It is found experimentally that suspensoids are almost always coagulated by sufficient addition of any soluble electrolyte. Furthermore, in many cases the concentration of electrolyte necessary to produce almost instantaneous coagulation is extremely small. However, the coagulating effect of various electrolytes can differ enormously. The effect is primarily on the rate of coagulation. Unfortunately, however, the technique of measuring coagulation rate and the general laws governing it have not been developed to a sufficient degree to make this an adequate method for studying the coagulating influence of electrolytes. Consequently, an empirical approach has been employed, in which one adds a measured quantity of the electrolyte in question to the suspension, with precautions to insure adequate mixing, and determines whether or not coagulation has occurred within a specified time. The experiment is repeated with varying amounts of added electrolyte, the minimum necessary to give precipitation under the conditions arbitrarily chosen being defined as the liminal concentration.*

TABLE III. GOLD SOL (NEGATIVELY CHARGED). LIMINAL
COAGULATING CONCENTRATIONS OF
DIFFERENT SALTS†

Electrolyte	γ millimols per liter	Electrolyte	γ millimols per liter
NaCl	24	$BaCl_2$	0.35
KNO_3	25	$Al_2(SO_4)_3/2$	0.009
$K_2SO_4/2$	23	$Ce(NO_3)_3$	0.003
HNO_3	10	—	—
$CaCl_2$	0.41	$UO_2(NO_3)_2$	2.8

Table III shows the liminal concentrations of various electrolytes for a gold sol. Inspection of this table brings out three

* In Freundlich's procedure [*Z. physik. Chem.* 44, 129 (1903)], a given volume of the sol is placed in a test tube, an equal volume of the electrolyte of known concentration added, pouring back and forth from test tube to test tube to secure thorough mixing. After standing for two hours, it is inspected. Two concentrations are found such that precipitation is just complete in one case and not in the other. The mean is taken as the liminal concentration.

† ELISSAFOFF, G. v., *Z. physik. Chem.* 79, 385 (1912).

important facts: In the first place, salts of similar type have approximately equal liminal concentrations, expressed as millimols or equivalents per unit volume. Thus, sodium chloride and potassium nitrate have practically the same liminal concentration, as compared with the far greater variations between salts of various types. Second, even among salts differing widely in type, the effect on liminal concentration is limited so long as the ion of charge opposite to that on the particle is unchanged. Thus, the precipitating effect of the nitrate and sulfate of potassium are about the same. Consequently, in view of the first point made, change in type of the salt has relatively little effect so long as the valence of the added contra-ion * remains unchanged. In the third place, inspection shows that change in valence of the contra-ion has an enormous influence on the precipitating effect of the electrolyte, the liminal concentration decreasing very greatly as the valence of the contra-ion goes up. Thus, while the merest trace of a tetravalent or trivalent electrolyte will coagulate a given sol, nearly 10,000 times as much monovalent ion may be required.

TABLE IV.   ARSENIOUS SULFIDE SOL (NEGATIVELY CHARGED).
LIMINAL COAGULATING CONCENTRATIONS
OF DIFFERENT SALTS†

Electrolyte	γ millimols per liter	Electrolyte	γ millimols per liter
KCl	49.5	$MgSO_4$	0.81
KNO_3	50	$SrCl_2$	0.635
HCl	30.8	$UO_2(NO_3)_2$	0.642
Strychnine nitrate	8.0	Benzidine nitrate	0.087
Toluidine sulfate	1.17	Quinine sulfate	0.24
Crystal violet	0.165	$AlCl_3$	0.093
$MgCl_2$	0.72	$Al(NO_3)_3$	0.095

Table IV shows similar results due to Freundlich in a sol of distinctly different type and Table V, also based on results of Freundlich, confirms the generalizations in a sol of opposite particle charge.

* Whatever the contra-ion originally present in the sol, addition of any electrolyte containing another ion of the same sign will result in immediate metathesis, so that the atter itself becomes a contra-ion.

† FREUNDLICH, H., Z. physik. Chem. **73**, 385 (1910).

Spectroscopic analyses made on coagulum and supernatant liquid by Annetts and Newman * afford additional facts of importance.   They prepared a negatively charged gold sol by

TABLE V.   FERRIC OXIDE SOL (POSITIVE).   LIMINAL
COAGULATING CONCENTRATIONS OF
DIFFERENT SALTS†

Electrolyte	Y millimols per liter	Electrolyte	Y millimols per liter
KCl	9.03	K_2SO_4	0.204
$BaCl_2/2$	9.64	$K_2Cr_2O_7$	0.194

the Bredig method, using sodium permanganate as the stabilizing electrolyte and magnesium chromate as coagulating agent. These salts were chosen since their ions contain heavy metals, easily detected and estimated by spectroscopic analysis.   When the gold sol was coagulated by freezing in liquid air and the gold particles separated in a centrifuge, it was found that 99% of the manganese was present in the coagulum and 95 to 98% of the sodium in the supernatant liquor.   When the gold sol was co-agulated with the liminal concentration of magnesium chromate, again all the manganese was found in the coagulum, while all the chromium remained in the supernatant liquor.   About 1/3 of the magnesium present was found in the coagulum, but this quantity became only 3 times as great as the concentration of the precipitating solution was increased 80 times.   Chromium was not found in the gold precipitate until 20 times the liminal concentration of magnesium chromate had been added.‡

The Mechanism of Precipitation with Electrolytes.

Early investigators believed that the coagulative powers, *i.e.*, the reciprocals of the liminal concentrations, stood in the ratio $1 : x : x^2$ for mono-, di-, and trivalent ions, x having a value of about 30 (Hardy-Schulze valence rule).   Different experimenters

---

* ANNETTS, M., and NEWMAN, L., *J. Phys. Chem.* **40**, 187 (1936) (The Williams and Wilkins Co.).

† FREUNDLICH, H., *Z. physik. Chem.* **44**, 129 (1903).

‡ These results assume even greater significance in the light of the discussion on pp. 129-130.

have found different ratios, *e.g.*,

Schulze [a]	1 : 49 : 810
Linder and Picton [b]	1 : 35 : 1023
Freundlich	1 : 104 : 810
Young and Neal [c]	1 : 39 : 875
Matsuno [d]	1 : 16 : 80 : 284

The ratio varies with the concentration of the suspensoid employed, and probably with other factors. The rule is completely upset if the precipitant forms an insoluble salt with the stabilizing ion of the suspensoid, as illustrated in Table VI for a mercuric sulfide sol, stabilized by hydrogen sulfide, keeping in mind that thallium, mercury, and copper all form insoluble sulfides.
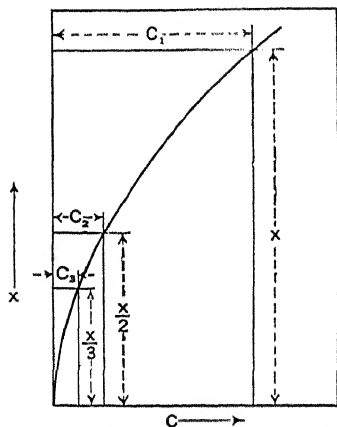
FIG. 6.   Adsorbability of Ions.*

Freundlich has attempted to explain the valence effect by assuming, first, that approximately equal numbers of charges must be adsorbed by the particle whether a mono-, di-, or trivalent coagulating ion is used and, second, that all the ions have approximately the same adsorption isotherm. If x mols of monovalent ion are required, $x/2$ of divalent and $x/3$ of trivalent ion will be necessary. Granting adsorption relations similar to those of Fig. 6, it is clear that very much smaller quantities of the di- and trivalent than of monovalent ions will be required.

The valence effect is to be anticipated from the known properties of the ions. Referring again to the gold particle in its sol (Fig. 5), since the hydrogen contra-ions in the neighboring solution are in active movement, a given ion will from time to time approach sufficiently closely to the chlorine ion adsorbed on the surface of the gold particle that the two can be considered

[a] Schulze, H., *J. prakt. Chem.*, **27**, 320 (1883).
[b] Linder, S. E., and Picton, H., *J. Chem. Soc.*, **67**, 63 (1895).
[c] Young, S. W., and Neal, R., *J. Phys. Chem.*, **21**, 14 (1917).
[d] Matsuno, K., *J. Tokyo Chem. Soc.*, **39**, 908 (1918).
* Freundlich, H., and Schucht, H., *Z. physik. Chem.* **85**, 641 (1913).

as undissociated, existing on the surface of the gold particle as an HCl molecule held by some type of force operative predominantly on the chlorine atom.   In view of the fact that the two

TABLE VI.   MERCURIC SULFIDE SOL *

Electrolyte	Liminal Concentration, millimols/l.	Electrolyte	Liminal Concentration
NaCl	13.0	HgCl$_2$	0.11
TlNO$_3$	0.14	Ce(NO$_3$)$_3$	0.056
HgNO$_3$	0.071	YCl$_3$	0.073
BaBr$_2$	0.68	CuSO$_4$	0.047
SrCl$_2$	0.88	Al$_2$(SO$_4$)$_3$	0.044

atoms carrying opposite charges are so close together, one would expect their mutual attractions to hold them there, so that the problem is to explain the fact that they do separate, the contra-ion wandering off again into the solution.† However, the fact that they again separate proves the existence of a disgregating force sufficient to overcome the energy corresponding to their mutual electrical attractions. Probably this force is intimately related, if not equivalent to the average kinetic translational energy of the molecules, whence one would expect its magnitude to be independent of the structure of the contra-ion in question. There will, therefore, be a dynamic equilibrium between the hydrogen ions approaching the gold particle (because of mutual electrical attraction) and those leaving it, owing to the disgregating action, whatever its nature. The hydrogen ions will be held at an average distance, at which the work done against the electrical attractions equals the energy of expulsion from the surface. The force of electrical attraction F, between two ions given by Coulomb's Law, is

$$F = \frac{\alpha\, n_1 n_2\, e^z}{S^2} \qquad (14)$$

where n_1, n_2 are the valences, S the distance between the two

* FREUNDLICH, H., and SCHUCHT, H., Z. physik. Chem. 85, 641 (1913).
† This question is, of course, the whole problem of the explanation of electrolytic dissociation.

ions, e the electronic charge, and α a constant. Granting that the energy of ionic disgregation which the free ion receives is constant, the work K, done in separating the ions from a distance S_0 to S, is given by

$$K = \int_{S_0}^{S} F\,ds = \alpha\,n_1 n_2\,e^2 \left(\frac{1}{S_0} - \frac{1}{S} \right). \tag{15}$$

In this expression S_0 is the distance between the adsorbed and contra-ion at their point of closest approach, *i.e.*, when they are both resident on the surface of the particle, in what is at least equivalent to an undissociated condition, and S is the average distance of the contra-ion in the solution, measured from the corresponding charge on the surface. $S\,S_0$ is the average distance traveled by the contra-ion away from the surface, and may be called the stand-off distance. The relation implies that no ions exist in the external solution other than the contra-ions, *i.e.*, that in the case in question substantially all the chlorine ions are adsorbed on the surface of the gold.

The relation just developed is only qualitative in character, but its implications can be brought out by making numerical assumptions. For example, while one does not know the value of S_0, the distance of closest approach of contra-ion to the surface charge, this unknown distance may arbitrarily be taken as unity. Assuming a univalent contra-ion, while one does not know S, there can be no doubt from the facts already quoted that S is considerably greater than S_0. For purpose of argument assume that S is four times S_0. The charge on the contra-ion n_1 is unity, while the charge on the particle n_2 is unknown. One is now in a position to calculate the value of $K/\alpha n_2 e^2 = 1 - 1/4 = 3/4$. If now the contra-ion be assumed trivalent, and any change in S_0 be neglected, one can compute its average distance from the surface charge S', by equating $K/\alpha n_2 e^2 = 3/4 = 3(1 - 1/S')$, whence $S' = 4/3$. Keeping in mind the fact that the average stand-off distance of the monovalent ion was $S - S_0 = 4 - 1 = 3$, whereas that of the trivalent ion is $S' - S_0 = 4/3 - 1 = 1/3$, it follows that the average distance of the monovalent ion from the surface charges on the particles is ninefold that of the trivalent ion.

The implications of this argument are indicated in parts (a) and (b) of Fig. 7. The structure of the particle surrounded by trivalent contra-ions is seen to be approaching that of the conventional Helmholtz electric double layer. The particle will be substantially neutral electrically with reference to any other particle outside the contra-ion stand-off distance. In contradistinction, the contra-ionic atmosphere of monovalent ions around a similar particle, shown in (a), is diffuse in character and large in radius. The repulsive effects of such a particle on a similar particle in its neighborhood would be large, reduced but little by the presence of the contra-ions. In consequence, in this case approaching particles of the suspensoid would begin to repel each other electrically while still far apart, by a force that would increase extremely rapidly as they approached. Possibility of contact is correspondingly reduced. In contrast, two similar particles surrounded by trivalent contra-ions [Fig. 7(b)] would not repel each other to any significant degree until they were very close together (until the shells of contra-ions were touching), and even then, on further approach the repulsive effect at any given distance between centers would be far less than in previous cases. Consequently, in a solution of monovalent contra-ions in which the charges on the particle are ample to give stability, replacement by trivalent contra-ions would diminish the stability to an extraordinary degree.

If electrolyte be added to a suspensoid which contains in the liquid only contra-ions, an additional effect will be superimposed on that just discussed. Thus, assume that the contra-ion is monovalent and the added electrolyte has the same ion [e.g., in Fig. 7(c) HCl is added to the sol]. The concentration of hydrogen ion will build up so that the conditions of Fig. 7(a) will be transformed into those of Fig. 7(c). Analysis of the

(a)

(b)

(c)

Fig. 7. Effect of Valency of the Contra-ions on the Ionic Atmosphere.

ionic relationships makes it clear that the presence of the added electrolyte will considerably reduce the distance at which the negative charge on the particle will be effective in repelling the similar charge on the neighboring particle, correspondingly reducing the stability. However, it will require considerable addition of such an electrolyte before the stability is reduced to a point comparable with that induced by trivalent contra-ions. Addition of even a small amount of an electrolyte having a polyvalent ion of the same sign as the monovalent contra-ion will result in what is equivalent to a metathetical displacement of the latter, converting the particle essentially into the structure of Fig. 7(b), with corresponding loss in stability. Despite the confusion resulting from complicated behavior in special cases, the general effect of valence on stability would seem to be a direct result of the electrical interactions of the ions, paralleling as it does the effect of valence on solubility.

It will be noted that all the electrolytes discussed up to this point have been inorganic in type. While organic salts follow the rules just developed in a qualitative way, the divergencies are far greater than those encountered with inorganic materials. In general, there is a tendency for the precipitating effect to exceed that indicated by the valence rule, particularly in the case of organic bases. The general character and magnitude of the effects will be seen from study of Table IV. The reasons for the abnormalities of these organic materials are not clear. Freundlich ascribes the effects to higher adsorption (see p. 91). Beyond doubt this explanation conforms to the known adsorption tendencies of the materials.

If one mix two suspensoid sols of opposite charge, the particles will attract each other and collisions cannot fail to occur. In the absence of secondary effects, one would expect collision to result in neutralization of charge, growth of particle size due to coagulation, and hence flocculation. In general, this co-precipitation is exactly what occurs. Mutual coagulation or equivalent effects are frequently employed for the coagulation of colloids (see pp. 383–384).

Since suspensoids are stabilized by an electric charge, one would anticipate that, despite the tendency to coagulate on

charge neutralization, if the neutralized charge could be replaced by one of opposite sign before coagulation could develop, a stable sol would again result (irregular series).   That this can be achieved in specific cases is shown by the results of Buxton and Teague in Table VII on the coagulation of a Bredig platinum sol by ferric chloride solution (see also p. 180).

TABLE VII.   COAGULATION OF A PLATINUM SOL BY $FeCl_3$*

Concn. of $FeCl_3$, millimols per liter	Degree of Coagulation	Particle Movement in Electric Field
0.0208–0.0557	No coagulation	To anode
0.0833–0.2222	Complete coagulation	No movement
0.3333–6.667	No coagulation	To cathode
16.33–666.7	Complete coagulation	No movement

The Physical Properties of Suspensoid Sols.

In general, the physical properties of suspensoid sols differ but little from those of the pure dispersing medium.   A principal cause of this is the impossibility of obtaining them in high concentration, a gold sol containing 5 gms. of gold per liter being exceptionally concentrated.   Most physical properties of a suspensoid are more or less additive, *i.e.*, the sum of the properties of the components.   Viscosity is an example of this.   Einstein has shown that the viscosity of a suspension of spherical particles should be given by

$$n_{sol} = n_{solvent}(1 + 2.5 \ V) \tag{16}$$

where V is the volume of the particles contained in 1 c.cm. of the sol.   This relation holds fairly well, although it is found experimentally that the viscosity usually does vary with the particle size and in many cases is not linear with the concentration.   In the derivation of the equation electrification of the particles was not taken into account; this may be responsible in some cases for deviations.

The Optical Examination of Colloidal Suspensions.

The properties of colloids depend to a large extent upon the size, shape, and structure of the individual particles and, there-

* BUXTON, B. H., and TEAGUE, O., *Z. physik. Chem.* 57, 64 (1906).

fore, the determination of these factors is very important. Moreover, the behavior of the colloid will depend upon the distribution of particle sizes in the sol. A sol containing particles of uniform size—a monodisperse system—shows properties differing from those of the same material containing particles of varying size—a polydisperse system. Furthermore, the properties of a polydisperse sol vary with particle-size distribution. Unfortunately, most of the methods available for study of particle shape and size are indirect in character.

The simplest way to determine the size of the particle in a sol is by direct microscopic observation. However, it is well known that two points cannot be "resolved," *i.e.*, distinguished in the microscope, when they are much closer together than the wave length of light. This limit of resolution in a microscope is given by the formula,

$$S = \frac{\lambda}{2n \sin \alpha} = \frac{\lambda}{2\text{N.A.}}$$

where S is the smallest distance which can be resolved using light of wave length λ, n is the index of refraction of the medium between the objective lens and the object, and α is half the angle at which the lens is seen from the particle. N.A., the product of the refractive index and half the angle of aperture of the lens, is called its numerical aperture. With visible light, of wave-length about 500 mμ, monobromnaphthalene, $n = 1.66$, as immersion liquid, and α as near to 90° as possible, the best microscopes are capable of resolving particles about 200 mμ in diameter. By the use of ultraviolet light of shorter wave-length, this resolution can be increased to include particles of 75 mμ diameter. Another factor limiting the use of the microscope is that for visibility the particles and dispersion medium must have distinctly different refractive indices. Consequently, little or no information can be gained on the size and shape of truly colloidal particles by direct visual observation in the ordinary microscope. The recently developed electron microscope gives resolutions hitherto unattainable.

It has long been known that a beam of light passed through a homogeneous fluid does not become visible to the eye, but if

dispersed particles are present the path may become clearly marked. This is known as the Tyndall effect. A familiar example is the path of the light ray from a projecting lantern to the screen, rendered visible by the cigarette smoke and dust particles in the air. If a beam of light be brought to a focus in a solution, it does not become visible when the solution is molecularly dispersed (and completely free from dust particles), but if the solution be colloidal, a bright cone, the Tyndall cone, is obtained. This is often used as a test for colloidal dispersions. The very fine particles scatter the light in all directions, even though too small to reflect the light in a regular manner. The scattered light is also plane-polarized, which allows the phenomenon to be distinguished from fluorescence, where the light is unpolarized.

This principle of causing the particles to emit light is embodied in the instrument called the ultramicroscope, devised in 1903 by Siedentopf and Zsigmondy.* In it the light scattered by the particles is observed through an ordinary microscope, appearing as brilliant spots against a dark background. The optical system used is shown in Fig. 8, where A represents a very intense source of light such as a carbon arc, projected on a fine slit B. The image of this slit, greatly reduced in size, is projected through the sol at C. In this way only a very thin section of the sol is illuminated. An ordinary microscope is placed at right angles to the illuminat-
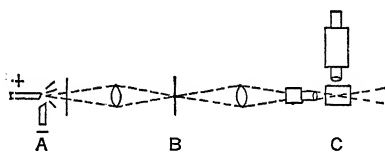
FIG. 8.   The Ultramicroscope.

ing beam to observe the particles. By this means it is possible to observe and count particles as small as 6 mμ, although of course their shape cannot be seen. The appearance of the spots of light in continual motion, owing to Brownian movement, does, however, give a clue to the shape of the particles. Spherical particles, since they are symmetrical, appear steadily bright, whereas rods and lamellae should tend to give rise to a twinkling effect, as they will scatter more light when their side of largest area is perpendicular to the illuminating beam.

* SIEDENTOPF, H., and ZSIGMONDY, R., *Ann. Phys.* (4) 10, 1 (1903).

Since the slit-ultramicroscope was devised, various other methods of illuminating the particle have been developed, in which the illuminating beam is also kept out of the eye of the observer. One of the most successful is the so-called cardioid dark-field condenser. The path of the light through this type of instrument is shown in Fig. 9. The concentration of the light on the particles is almost ideal, as all the light in the illuminating beam is focused at a point. Figure 10 shows the type of cell used, made of fused quartz (since glass is slightly fluorescent). In the lower part there is a ring-shaped groove, the quartz inside which is polished down a few thousandths of a centimeter, so that by placing a drop of the sol on the lower plate and covering it with the cover glass, a thin layer of solution enclosed between the quartz surfaces is obtained. This type of instrument is easy to manipulate and exceedingly convenient. Like the microscope, the ultramicroscope is useless if the refractive indices of particle and dispersion medium are nearly the same.

FIG. 9. The Cardioid Dark-Field Condenser.

FIG. 10. Cell for Use with Cardioid Condenser.

In some cases information on the shape of the particles can also be obtained by studying the scattering of light from the particles when the sol is under the influence of external forces. As already mentioned, more light will be scattered when the large surface of the particle is perpendicular to the incident light beam. Thus, if a sol containing rod-shaped particles be caused to flow through a cell, the rods become oriented in the direction of flow and the intensity of the light scattered in a perpendicular direction will be much greater than that scattered if the sol is viewed in the direction of flow.

Double refraction provides another means of obtaining information on the shape of the particles. If a sol is placed in a

hollow prism and a beam of polarized light passed through it, this light ray will behave in the normal way, as shown in Fig. 11(a). If the sol be caused to flow through the prism, in some cases its behavior remains unchanged, while in others a secondary extraordinary refracted ray is produced [Fig. 11(b)]. This is known as "stream double refraction." Double refraction can at times be obtained by placing the sol in a magnetic field, causing "magnetic double refraction." In any event

(a)   (b)

FIG. 11. Stream Double Refraction.

double refraction must be caused by orientation of the particles. By studying these phenomena valuable information on particle shape may be secured.

Methods of Purification.

An excess of soluble, low molecular weight material, often ionic, is generally present in the preparation of colloidal solu-- tions. This may be high enough in concentration to cause ultimate flocculation of the resulting sol and in most cases it is desirable to reduce the amount of the impurities as much as possible. The important methods of purification so far devised are based on the enormous size of the colloidal particle (5 mμ to 0.5 μ) compared with that of the molecularly or ionically dissolved components of the solution (about 0.1 mμ), and take advantage of the selective permeability of suitable membranes. The three most common methods are dialysis, electrodialysis, and ultrafiltration.

Due to its great size, the velocity of diffusion of the colloidal particle, which varies approximately inversely as the square root of the size, is very low. Moreover, many natural and artificial membranes can be obtained, the pores of which are smaller than the colloidal particle but very much larger than any molecularly dissolved material. Therefore, by separating a colloidal solution from a quantity of its pure dispersion medium by means of such a membrane, the molecularly dissolved constituents will diffuse into the region of lower concentration at an appreciable rate, due both to their high diffusion velocity and to the fact that the membrane offers no appreciable resistance. The par-

ticles of colloidal size, however, move very slowly and cannot penetrate the membrane, so that they remain in the solution and a separation can easily be effected. It was in this way that Graham in his classical researches was led to distinguish colloidal and crystalloidal substances. The former, albumin, gelatin, etc., when dissolved in water do not pass through a parchment membrane when submitted to dialysis, while the latter, sugar, salt, etc., readily pass through. For this reason Graham thought that he had discovered two distinct states of matter.

A suitable apparatus for purifying colloidal sols by dialysis is shown in Fig. 12. The impure sol is placed in the compartment A, separated from the compartment B, containing pure solvent, by a semipermeable diaphragm C. The diaphragm can be made of such materials as parchment, collodion, or Cellophane. The efficiency of the apparatus is much increased if a stream of pure solvent be passed through B, so that the soluble constituents are removed as rapidly as they diffuse. Efficiency of dialysis depends, in general, upon the type of constituent to be removed, the permeability of the membrane, the dimensions of the apparatus, and the temperature at which dialysis is carried out.
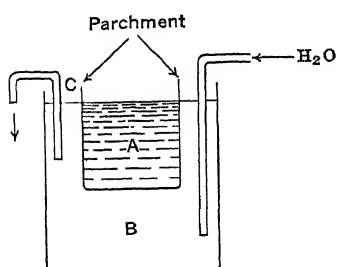
FIG. 12.   Dialysis Apparatus.

Electrodialysis combines dialysis with electrolysis. The passage of ionic impurities through the membrane is hastened by applying a potential difference across the dialyzer. A cheap electrodialyzer, easily made from old storage batteries, is shown in
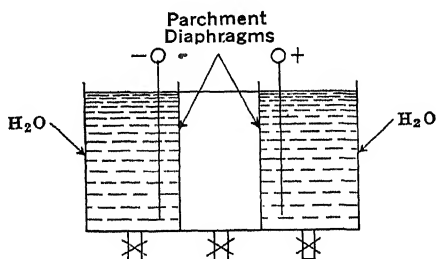
FIG. 13.   Electrodialysis Apparatus.

Fig. 13. Electrodes are placed outside the diaphragms and water is circulated as before through the outer compartments. By this means, the removal of any ionic constituent in the solution is greatly facilitated, the speed

at which the process occurs being increased and ultimate purification of the solution improved.

In ultrafiltration the solution is forced under pressure through a membrane permeable only to the constituent of the solution that it is desired to remove.   It is thus a combination of dialysis with ordinary filtration.   The part of the solution remaining behind becomes more concentrated in the colloidal constituent, but the removal of foreign electrolyte is never as complete as in simple dialysis.   On account of the more difficult experimental technique, ultrafiltration is not as extensively employed as the other two methods for purifying colloidal solutions.

Viscosity Characteristics of Suspensions.

The dispersions discussed up to this point have been predominantly dilute.   As concentration of the dispersed material is increased, the properties of the dispersion change progressively, perhaps the most important change being in the viscosity. With increase in concentration, the Einstein equation always breaks down, apparently rarely holding beyond about 8 to 10 volume per cent and frequently failing far below this limit.   The viscosities of concentrated suspensions exceed those indicated by the equation.   When concentration becomes sufficiently high, abnormalities in flow also develop, in the sense that the frictional resistance to shear is no longer

Fig. 14.   Viscosity Characteristics of Some Aqueous Suspensions.*

proportional to the velocity gradient (cf. p. 220).   This is brought out in Fig. 14, in which each curve indicates the flow characteristics

* THOMPSON, W. I., *Mass. Inst. Tech. Thesis* (1935).

of an aqueous suspension of the volumetric concentration shown, the measurements being made in a concentric cylinder viscosimeter. Ground silica, even when present to the extent of 62 volume per cent, shows normal flow behavior, the shearing force still being proportional to the rate of shear. In contrast, calcite, mica, and china clay in concentrations ranging from 32 to 40% are all abnormal. For these materials the relation between shearing stress and velocity gradient is nearly linear over a considerable range. Furthermore, when the lines are extrapolated they all cut the shearing stress axis at positive values. This would indicate that a finite shearing stress, the so-called yield value or yield point, is necessary to initiate flow.* Because a true fluid is completely incapable of permanently sustaining a shearing stress, the behavior just described can be explained only on the assumption of a structure existing in the dispersions, which must consist of some sort of a substantially continuous, solid framework (see pp. 219–229).

Where abnormal flow is encountered, it is clear that viscosity becomes a function of rate of shear, so that redefinition of terms is necessary. While terminology is not standardized, in this text viscosity will be defined as the ratio of the total shearing force per unit area to the corresponding rate of shear (p. 19). Curve I of Fig. 15 shows the velocity gradient-shearing force curve for an ordinary liquid; clearly, its viscosity corresponds to the ratio of the line AB to BC. Its fluidity is the reciprocal of this quantity. The flow characteristics of many suspensions may be represented by curves similar to II of Fig. 15 (vide infra); clearly, viscosity as defined above is dependent here upon the shearing force. For the point D, $\mu = AE/DE$. The slope

FIG. 15. Shearing Force *vs.* Velocity Gradient.

---

* In many cases the curve of shearing stress against rate of shear bends rapidly as the stress axis is approached. Sometimes it comes into the origin with finite slope, but often with negligible slope. While these facts must never be forgotten, it is convenient to extrapolate the lines here shown to the intercept as though they remained straight.

of the curve is defined, therefore, as the mobility and, as stated, the shearing force necessary to initiate flow is the yield value.   Providing the relation of Fig. 15 is linear, mobility is constant and is numerically equal to the viscosity at infinite rate of shear.   Where curve II of Fig. 15 exhibits curvature, both mobility and viscosity vary with rate of shear.

The effect of a given volumetric concentration on flow characteristics is greatly influenced by the shape of the suspended particles.   This is indicated in Fig. 14, in which the materials exhibiting yield values at low volumetric concentration are all well recognized to be platy in structure.   In general, the greater the ratio of one or two of the dimensions of a particle to the third, *i.e.*, the greater its asymmetry, the lower the concentration required to develop abnormal flow.   Particle agglomeration is also a factor which may profoundly influence flow behavior. In many cases, this is probably due to the tendency of agglomeration to build up particles of asymmetric shape but relatively stable structure.   In this sense, its effect is an illustration of the influence of particle shape.

The decrease in viscosity with shearing rate from an infinite value at zero rate of shear must be caused by the influence of flow on the structure of a suspension.   The importance of this must be discussed in detail elsewhere (see p. 245), the extraordinary complexity of the relationships which may be encountered also being emphasized.

BIBLIOGRAPHY

BURTON, E. F., *The Physical Properties of Colloidal Solutions*, Ed. 3, Longmans, London (1938).

FREUNDLICH, H., "The Optical Anisotropy of Colloidal Solutions," *Science Progress* 30, 218 (1935).

FREUNDLICH, H., *Kapillarchemie*, Ed. 4, Akad. Verlag. m.b.h., Leipzig (1930).

KRUYT, H. R., *Colloids*, Ed. 2, Wiley (1930).

PERRIN, J. B., *Atomistique*, Hermann, Q., Paris (1935).

SVEDBERG, T., *Colloid Chemistry*, Ed. 2, Chem. Cat. Co. (1928).

Chapter VII

Amorphous Solids

One class of colloidal solutions, the suspensoids, was considered in the previous chapter. However, the larger number of colloidal solutions are of another type, industrially far more important and of strikingly different properties. These are most frequently produced by direct solution of amorphous solids in appropriate solvents, and, in order fully to appreciate their characteristics, it is essential to develop the clearest possible picture of the chemical structure of the amorphous materials from which they are prepared. The application of the classical methods of determining the structure of chemical compounds to natural amorphous materials, such as rubber, cellulose, proteins, and the like, was early recognized as inadequate. The materials were difficult to purify; ordinary osmotic methods for determination of molecular weight broke down because the compounds gave values too high for accurate measurement; furthermore, no methods of chemical synthesis could be developed. The progress of recent years in solving these problems has been extraordinary; electrodialysis, the centrifuge, etc., have improved methods of purification; the ultracentrifuge and viscosity studies give dependable determinations of molecular weight; finally, direct and relatively simple syntheses, not, it is true, of the natural products themselves, but of products of remarkably similar properties, have been realized. These results have opened a new chapter of extraordinary promise in the study of amorphous materials.

The Linear Polyesters.

Consider the possibilities of chemical interaction between ethylene glycol and succinic acid. The simplest reaction which

can occur is that between one molecule of each of the compounds,

$$HO.CH_2.CH_2.OH + HOOC.CH_2.CH_2.COOH \rightarrow$$
$$HO.CH_2.CH_2.O.CO.CH_2.CH_2.COOH + H_2O.$$

The product, the mono-monoester, still contains a residual hydroxyl and a corresponding carboxyl group. It is, therefore, capable of reaction either with more glycol, with more succinic acid, or with itself. In the last case it would form the lactone-type compound,

$$\overline{O.(CH_2)_2.O.CO.(CH_2)_2.CO}$$

a neutral body no longer possessing either hydroxyl or carboxyl groups and consequently incapable of further esterification.* In either of the two former cases, esterification would produce a chain compound of increased molecular weight, always possessing at the end of the molecule reactive groups, hydroxyl or carboxyl as the case may be, making it capable by suitable esterification of still further increase in chain length and molecular weight.† The fundamental unit in the chain is

$$—O—(CH_2)_2—O—CO(CH_2)_2CO—O—$$
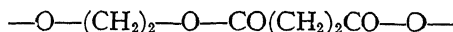
which may be repeated an indefinite number of times, but there are always hydrogen atoms combined with the terminal oxygen atoms.

It is found experimentally that, if approximately equi-molecular quantities of ethylene glycol and succinic acid are heated together, they react rapidly; the reaction can be com-pleted and the water removed by raising the temperature to 200 or 225° C., under an absolute pressure of approximately 1 mm. of mercury. The resulting product is a viscous liquid, which solidifies on cooling to a waxy mass, readily soluble in benzene or chloroform, from which it may deposit well-formed, microscopic crystals. It is a mixture, but can be fractionated by suitable techniques. By using a small excess of glycol in the initial mixture, one can obtain a crystalline product with a quite

---

* This can result not only from interaction of the hydroxyl and carboxyl groups of a single molecule but of any molecule of this type.

† Conceivably, this process might be stopped at any stage by lactone-type interaction of the alcoholic and acidic groups on opposite ends of the chain.

well-defined melting point of 102° C. and a molecular weight, determined in various solvents by both freezing- and boiling-point methods, of approximately 3000. The product is neutral and gives (somewhat sluggishly) the reactions of the hydroxyl group. Thus, it reacts with *p*-brombenzoic anhydride to give a derivative, the bromine content of which corresponds with that required by the formula,

$$HO.[(CH_2)_2{-}O{-}CO(CH_2)_2CO{-}O]_{22}{-}(CH_2)_2.OH$$
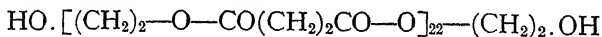
equivalent to a molecular weight of the original product of 3230. If, on the other hand, one employs a small initial excess of succinic acid, the product is acidic with a melting point of 98° C. Carothers determined the molecular weight of such a product by (*a*) cryoscopic methods, (*b*) titration with alkali, and (*c*) determination of the sodium content of its salt, finding 3110, 3400, and 3740, respectively. The formation of compounds of such high molecular weight forces the conclusion that in this reaction, long chain, linear polyesters are formed, the terminal groups being determined predominantly by the initial proportions of the reactants. Length of chain can be controlled partly by the proportions of the reactants and by choice of the operating conditions.* Furthermore, the physical properties of the esters depend significantly upon the length of chain. Members of molecular weight below 5000 are soft and dissolve readily in solvents, but above 10,000 they are tougher, harder, and swell in the process of dissolving (pp. 236, 305).

* No evidence of lactone formation is normally observed with ethylene glycol and succinic acid, although the lactone can be prepared. With oxalic acid, the lactone is formed smoothly, presumably because it consists of a six- rather than an eight-membered ring.

$$HO{\cdot}CH_2{\cdot}CH_2{\cdot}OH + (COOH)_2 \longrightarrow$$
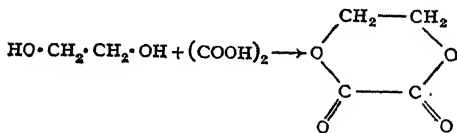
Even in this case, however, the lactone tends to transform into a high molecular weight, linear polyester.

Cyclic polyesters with more than five or six atoms in the ring are rarely formed directly but can result from drastic treatment of the linear polyesters. This is believed to be due to the spatial configuration; models indicate that the chance of active terminal groups approaching sufficiently closely for reaction is almost negligible when they are separated by more than four atoms.

Linear polyesters may also be formed by the self-esterification of hydroxy acids, *i.e.*, compounds containing both hydroxyl and carboxyl groups in the same molecule, where the hydroxyl and carboxyl groups are separated by more than four carbon atoms to minimize lactone formation. The ω-hydroxydecanoic acids, $HO.(CH_2)_9.COOH$, investigated extensively by Carothers and other workers, are highly instructive. These acids show no tendency towards lactone formation but, on being heated to 150 to 250° C. at reduced pressure, react to produce viscous liquids which solidify to waxy solids. In appearance and physical properties the esters resemble those produced from ethylene glycol and succinic acid. By reducing the pressure to 10^{-4} mm. mercury and raising the temperature to 200 to 250° C., molecular weights as high as 25,000 have been obtained.

The most important characteristics of these polymers are their physical properties and the influence of molecular weight * upon them, as summarized in Table I. The lower members of the series behave like ordinary organic compounds. However, as the molecular weight increases, not only does the viscosity of the resulting solutions go up abnormally, but two other significant phenomena are observed. In the first place, the materials of high molecular weight dissolve relatively slowly, and solution itself is preceded by considerable swelling of the solid, caused by imbibition of solvent. In the second place, if one touches the liquid surface of the molten ester of high molecular weight with a rod and draws it away slowly, the liquid is spun out into long, delicate fibers. Similar fibers can be produced by extruding a solution of the same polyester through a fine orifice into a current of warm air into which the solvent can evaporate (cf. the manufacture of acetate rayon, p. 373). These filaments as first formed are opaque, brittle, and give an X-ray diagram (Fig. 1) which is characteristic of non-oriented or only slightly oriented crystallites. If, however, the filaments are stretched (*i.e.*, cold-drawn), they can be permanently elongated, and then become relatively transparent and mechanically much stronger.

* There is no reason to consider any specific product homogeneous in the sense that all molecules in it are identical in size. Any molecular weights measured are averages. However, fractions can be formed in which the variations in molecular weight are probably not excessive.

At the same time the X-ray diffraction pattern changes (Fig. 2), indicating considerable orientation of the crystallites in the direction of stretch. As indicated in Table I, the higher the molecular weight of the material, the more striking are these unusual phenomena.
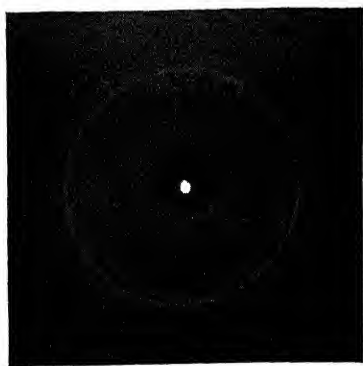
Fig. 1.　X-Ray Diagram of Linear Polyester.*
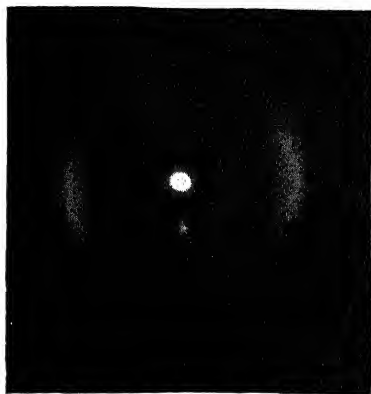
Fig. 2.　X-Ray Diagram of Stretched Linear Polyester.*

The probable explanation of fiber formation is a certain parallel orientation of the ultimate molecules in the particle. Initially, this is only incipient, but elongation by drawing accentuates the parallel arrangement, the molecules being combed into position by slip. Lowering of temperature increases intermolecular friction (by accentuating the mutual attractions of the molecules, because of reduced distance due to increased density, and lessening the disgregating forces of thermal agitation; see p. 12) and hence interferes with drawing, but greatly increases fiber strength, provided the effect is not sufficient to embrittle. Swelling is due to the gradual opening up of the molecules in the solid from each other, along at least a portion of their length, sufficiently to allow penetration of solvent into the solid before the individual molecules are completely separated or loosened from the solid, as is necessary for dispersion and solution.

The parallelism in physical properties between these polyesters, which, in view of the mechanism of synthesis, are evi-

* CAROTHERS, W. H., and HILL, J. W., *J. Am. Chem. Soc.* 54, 1557 (1932).

TABLE I. PROPERTIES OF ω-HYDROXY-DECANOIC ACID POLYESTERS *

Av. Mol. Wt. (titration)	Calcd. Length of Molecule	Melting Point	Density, 25°/4°	Fiber Character	Mol. Wt. (other methods)	Tensile Strength, kg./mm.²
780	60 Å	66–67° C.	1.0957	Absent	930 (boiling benzene)	
1,720	123	72–74	1.0935	"	1,630 (boiling benzene)	
3,190	188	74–75	1.0877	"		
4,170	313	74–76	1.0814	"		
5,670	440	73–75	1.0751	Very short fibers Do not cold draw		
7,330	570	74–75	1.0715	Long fibers Do not cold draw		Very weak
9,330	730	75–76	1.0668	Long fibers, which cold draw		Very weak
16,900	1,320	77–78	1.0627	Easily spins, and cold draws		13.1
20,700	1,610	77–78	1.0632	Spins with difficulty, but easily cold draws	26,700 (ultracentrifuge)	12.3
25,200	1,970	75–80	1.0621	Spins above 210° C. and cold draws		7.0

* Carothers, W. H., and Van Natta, F. J., *J. Am. Chem. Soc.* 55, 4714 (1933).

dently long-chain molecules, and many of the industrially most important organic materials, such as cellulose, wool, silk, and rubber, justifies the inference that the natural products likewise owe their properties to similar chemical structure.*

Linear polymers † of high molecular weight can be formed by other reactions than esterification. Thus, the reaction of glycol with acetaldehyde gives *polyacetals*, water being eliminated between two hydroxyl groups and the carbonyl group. Products formed by interaction of amino and carboxylic radicals are being introduced industrially (see p. 377).

$$2\ HO.CH_2.CH_2.OH + CH_3.CHO \rightarrow$$
$$HO.(CH_2)_2.O.CH(CH_3).O.(CH_2)_2.OH + H_2O$$
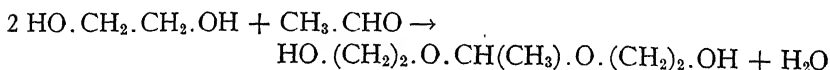
This acetal is capable of further reaction with acetaldehyde, giving structures of the type:

$$HO\text{—}R\text{—}O\text{—}\underset{\underset{CH_3}{|}}{CH}\text{—}O\text{—}R\text{—}O\text{—}\underset{\underset{CH_3}{|}}{CH}\text{—}O\text{—}R\text{—}O\text{—}\underset{\underset{CH_3}{|}}{CH}\text{—}O\text{—}$$
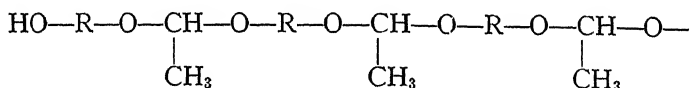
In this, R represents the glycol residue. *Polyanhydrides* result where water is eliminated between carboxyl groups of dibasic acids, which contain eight or more carbon atoms, producing a linear compound,

$$HOOC\text{—}R\text{—}CO\text{—}O\text{—}CO\text{—}R\text{—}CO\text{—}O\text{—}CO\text{—}R\text{—}CO\text{—}$$
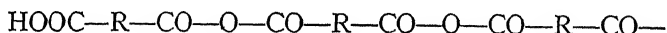
Carothers has termed this type of reaction, giving rise to a long-chain compound in which two groups of each molecule are active, bifunctional. It may be represented by the general equation:

$$x - R - x\ +\ y\cdot\cdot R' - y\ =\ -R' - z - R - z - R'\ +$$
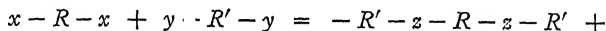
where x and y represent the functional groups, z the combination of x and y, a the product eliminated by the reaction, and R and R' the inert parts of the reacting molecules. In general, a bifunctional reaction may well result

---

* This opinion has been held for several decades but, despite the indications which justified it, it is only with the accumulation of further evidence of the type here described that in recent years chemists in general have begun to accept this interpretation.

† Strictly speaking, the word "polymer" should be limited to reaction products having the same percentage composition as the original reactant, or even to products of direct molecular addition. Thus, benzene may be described as a polymer of acetylene. Recently, however, the word has been extended to cover any high molecular weight material produced from simple molecules, whether by condensation, i.e., by elimination of water, etc., or by true polymerization.

in a linear polymer as above or, less frequently, in a closed ring:

$$x - R - x \; + \; y - R' - y \; \rightarrow \; \underset{R'}{\overset{R}{\underset{\diagdown}{\diagup z \quad z \diagdown}}} \quad \text{or} \quad \underset{R' - z - R - z - R'}{\overset{R - z - R' - z - R}{\diagup \qquad\qquad \diagdown \atop \diagdown \qquad\qquad \diagup}}$$
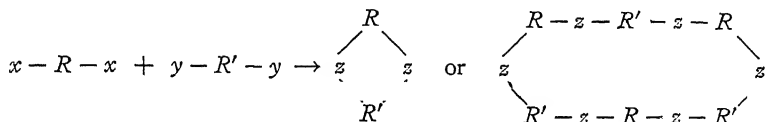
The size of the ring or length of the chain is determined by reaction conditions. When a molecule contains more than two functional groups, the reaction becomes much more complicated, with the possibility of forming other types of structures (pp. 155–158).

Styrene and Isobutylene.

Compounds of high molecular weight can also be formed by true polymerization, in contrast to condensation, particularly where two or more unsaturated molecules of the same kind combine to produce a resultant, more saturated molecule. Thus, ethylene under favorable conditions can polymerize to give butylene

$$2 \; CH_2{=}CH_2 \rightarrow CH_3 . CH_2 . CH{=}CH_2,$$

the total unsaturation being reduced from two to one double bond. Since the product is also unsaturated, further reaction can occur, building up high molecular weight compounds of the type $CH_3{-}(CH_2)_x{-}CH{=}CH_2$. The polymerization of ethylene is difficult but substituted ethylenes frequently polymerize readily to give good yields of fairly homogeneous products. The following facts regarding polymerization of phenyl ethylene or styrene, $C_6H_5CH : CH_2$, seem well established.

(1) When styrene is heated, the liquid gradually becomes viscous and sets to a jelly-like mass. Light and catalysts, such as oxygen and stannic chloride, are particularly effective in accelerating the reaction, while anti-oxidants, *e.g.*, hydroquinone, inhibit it almost completely

(2) The molecular weight of the polymers varies with conditions of polymerization from 1000 to 600,000. Each product is probably non-homogeneous, capable of fractionation by solution and by fractional precipitation.

(3) The polymers give no detectable bromine absorption, thus showing no evidence of unsaturation.

(4) Hydrogenation involves addition of six hydrogen atoms for each benzene ring originally present, the treatment being without sensible effect on the apparent molecular weight of the polymer.

(5) Heated above 300° C., polystyrene breaks down to give large yields of styrene. These facts make probable the following formula for polystyrene:

$$\underset{|}{C_6H_5} \qquad \underset{|}{C_6H_5} \qquad \underset{|}{C_6H_5} \qquad \underset{|}{C_6H_5}$$
$$-CH-CH_2-CH-CH_2-CH-CH_2-CH-CH_2-$$

or perhaps

$$\underset{|}{C_6H_5}\ \underset{|}{C_6H_5} \qquad\qquad \underset{|}{C_6H_5}\ \underset{|}{C_6H_5}$$
$$-CH_2-CH-CH-CH_2-CH_2-CH-CH-CH_2-.$$
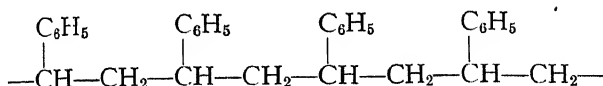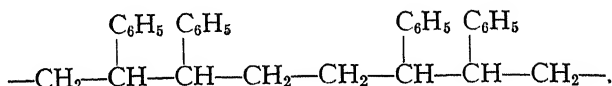
The mechanism by which unsaturated compounds polymerize is in many cases far from clear and the character of the terminal groups unknown. In a compound such as polystyrene, containing many hundred styrene residues, the presence of a single unsaturated terminal group is exceedingly difficult to detect. Nevertheless, the reactivity of the lower polystyrenes to more monostyrene, and the fact that unsaturation has been detected in the closely related indene polymers:

$$\underset{|}{C_6H_4}-\underset{|}{CH_2} \left[\underset{|}{C_6H_4}-\underset{|}{CH_2}\right] \underset{|}{C_6H_4}-\underset{|}{CH_2}$$
$$-CH_2-CH-\left[CH\ -CH-\right]_xC\ =\ CH$$
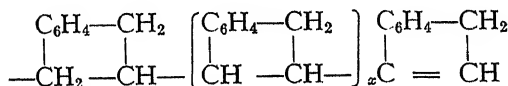
give weight to the assumption of a long-chain structure terminating with an unsaturated grouping:

$$\underset{|}{C_6H_5} \qquad \left[\underset{|}{C_6H_5}\right] \qquad \underset{|}{C_6H_5}$$
$$CH_2-CH_2-\left[CH-CH_2-\right]_xC=CH_2$$
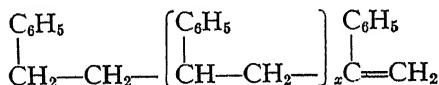
These polymers also vary in physical properties with molecular weight. Thus, low molecular weight polystyrene (ca. 4000) is an easily pulverizable powder, soluble in benzene without swelling, and with a liquefaction range of 125 to 170° C. In contrast, a polymer of molecular weight 100,000 or over melts from 250 to 300° C. and the tough, elastic mass swells strongly before dispersion in benzene.

Isobutylene, $CH_2=C(CH_3)_2$, cooled to a low temperature with a catalyst such as boron fluoride, readily gives linear polymers with no side reactions. The reaction mixture may set to an almost solid mass. By suitable adjustment of reaction conditions polymers with average molecular weights of up to 300,000 can be obtained. The lower polymers (molecular weight 3000 to 20,000) are tacky, very viscous liquids, while the higher ones are extensible solids resembling rubber in superficial appearance. The polymers are soluble, the higher ones swelling before going into solution, and having a much greater solubility than a paraffin of corresponding molecular weight. On heating, the polymers revert to isobutylene, which distils off. The polymerization may be represented by the equation

$(n + 2)C(CH_3)_2=CH_2 \rightarrow (CH_3)_2CHCH_2-$
$$[CH_2CH(CH_3)CH_2-]_nCH_2C(CH_3)=CH_2.*$$
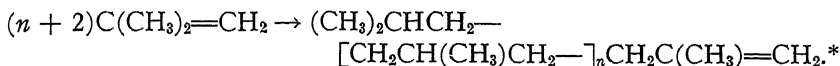
Presence of n-butylene in the isobutylene yields for given conditions a polymer of lower average molecular weight. Presumably, although a n-butylene molecule can attach itself to an isobutylene chain, the resulting terminal double bond is too inactive for more chain growth.

Other substituted ethylenes exhibit similar behavior. Thus, vinyl acetate, $CH_2=CH.O.CO.CH_3$, polymerizes to give a product (p. 479) having the probable structure:

$$-CH_2-CH-CH_2-CH-CH_2-CH$$
$$O.CO.CH_3 \ \ \dot{O}.CO.CH_3 \ \ \dot{O}.CO.CH_3.$$
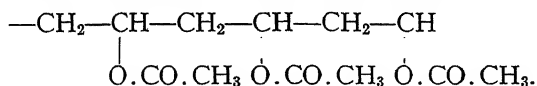
There are natural materials of outstanding importance which, while the chemist has not yet succeeded in synthesizing them, exhibit physical properties and chemical behavior leading to the conclusion that they are linear chain polymers of the type just discussed. Among these, cellulose and rubber deserve special mention.

Cellulose.

Cellulose is probably the most important naturally occurring amorphous solid of high molecular weight. It is widely dis-

---

* It should be noted that there is a double bond at the end of the chain, though the molecule is otherwise saturated.

tributed in nature (Table I, Chapter XXI) but has properties which vary relatively little with its source. It is extraordinarily inert, infusible below its temperature of thermal decomposition, insoluble in ordinary solvents, and, in general, dissolves only in liquids containing chemically active solutes. Its ultimate analysis indicates an empirical formula $C_6H_{10}O_5$.

The close chemical relationship between cellulose and the sugars is shown by its hydrolysis in solutions of high hydrogen-ion concentration to d-glucose as final product.* The hydrolysis is step-wise, and intermediate stages which have been recognized are cellodextrins of varying complexity, a cellotetrose, a cellotriose and, most important of all, cellobiose, $C_{12}H_{22}O_{11}$. Unlike the intermediate products and glucose, to which it hydrolyzes, cellulose has very little tendency to reduce Fehling's or similar solutions, $i.e.$, it has a low copper number.† On the other hand, no matter how carefully natural cellulose is purified, the product is never completely free of reducing power, although the purest specimens have copper numbers as low as 0.05, in contrast with the value of 300 for d-glucose. Moreover, if cellulose is submitted to conditions tending toward hydrolysis, its copper number always increases. Thus, dilute weak acids at low temperatures will cause slight rises in copper number, although the conditions are such that no trace of ultimate hydrolysis to glucose is detectable. That a fundamental change in the structure of the cellulose accompanies these small increases in copper number is confirmed by the simultaneously occurring changes in physical properties. Thus, the exposure to cold, dilute, weak acid referred to will cause a drop in the tensile strength of a cellulose fabric, or even of an individual cellulose fiber tested under comparable conditions. These facts point inescapably to the conclusion that cellulose is a condensation product of d-glucose, formed by elimination of water, by a chemical mechanism which eliminates largely, but not quite completely, the aldehyde groups of the sugar.

---

* Yields of glucose as high as 95% of the theoretical have been obtained.

† Fehling's solution contains copper sulfate, sodium potassium tartrate (Rochelle salt), and caustic soda, with the copper in a complex ion. The copper number is defined as the weight (in grams) of copper reduced, under carefully specified conditions, from the cupric to the cuprous state by 100 gms. of the material under test.

Cellulose reacts readily with many acids to form esters, and yields ethers by alkylation. Under mild conditions of conversion the shape and apparent structure of the original fiber may be retained unimpaired. Moreover, where it is possible to remove the reactive groups under similar conditions, the cellulose fiber is restored with its properties substantially unchanged, except that its tensile strength has usually dropped a few per cent. The esters and ethers are soluble in organic solvents. They show high molecular weights, as indicated by negligible osmotic effect; as measured by ultracentrifuge technique (p. 107), values above 50,000 are obtained. When the reactions are carried to completion under conditions avoiding obvious decomposition of the cellulose, three monovalent radicals react with each glucose residue, indicating the presence of three hydroxyl groups, in contradistinction to the five of glucose.

Any satisfactory structural formula for cellulose must account for its high molecular weight, its progressive, step-wise conversion by hydrolysis to cellobiose and finally, glucose, and the facts that it contains only three hydroxyl groups for each glucose residue and retains traces of aldehyde groups, of which the quantity increases progressively with hydrolysis. The problem has a long history, but it is now agreed that cellulose is made up of cellobiose residues resulting from the elimination of water between two hydroxyl groups of the β form of d-glucose.* In
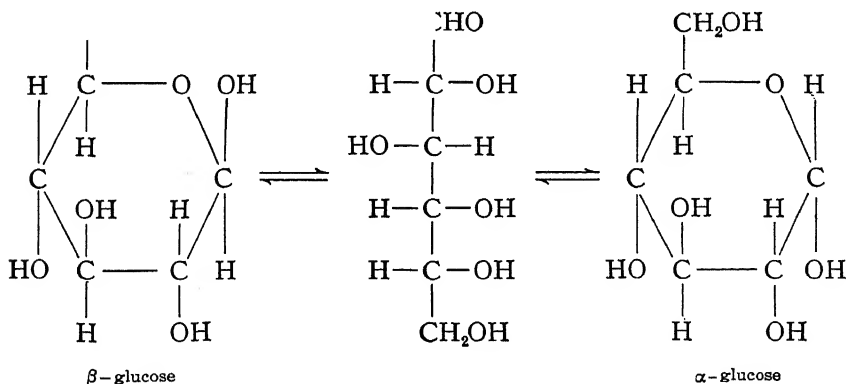
β – glucose α – glucose

* Haworth and his collaborators have located the hydroxyl groups on the ultimate glucose residue in the cellulose molecule, e.g., by methylating cellulose, hydrolyzing the product under conditions avoiding disturbance of the methoxy groups, and identifying their location on the products of the hydrolysis.

this condensation, water is eliminated from the hydroxyl group on the end carbon atom of the β-glucose molecule, the location of the original aldehyde group, and a γ-hydroxyl group of a neighboring β-glucose molecule, giving the structure:
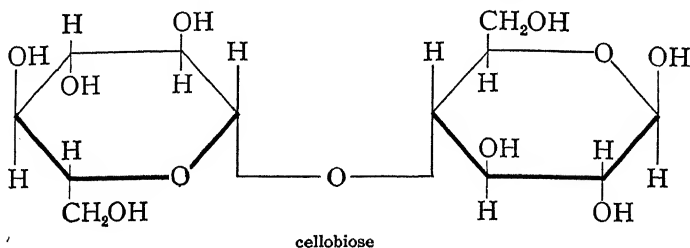
<div align="center">cellobiose</div>

It is clear that a condensation of this sort can continue, building up the molecule by the addition of successive glucose radicals to form a chain of indefinite length and molecular weight. The
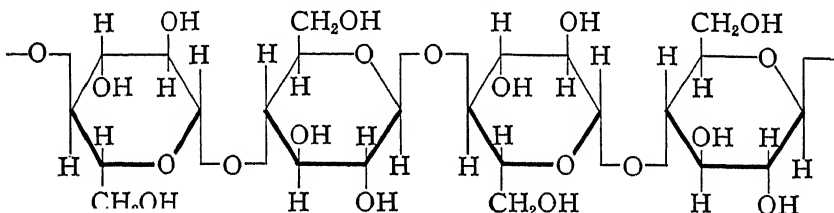
isolation of cellobiose from the products of hydrolysis of cellulose indicates that this is the structure, rather than a similar grouping in which α-glucose is involved. Actually, this latter configuration is found in starch (p. 310).
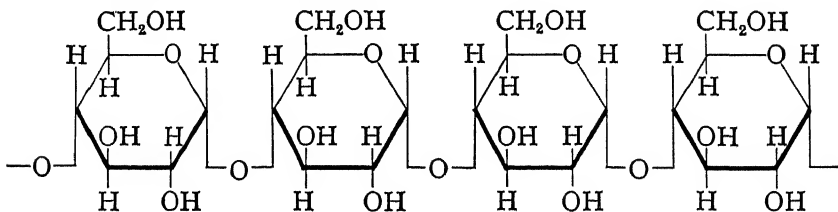
Such a linear, chain-polymer structure for cellulose affords an explanation of all the above facts. Its physical inertness corresponds to its high molecular weight. The aldehyde group no longer remains as such in the molecule, except in potential form on the final glucose residue of the chain. Hence, in the absence of hydrolysis, cellulose must behave chemically as an alcohol

with three hydroxyl groups to each glucose residue. In the long-chain structure, the ether linkages are, by and large, alike. Hydrolysis should, therefore, tend to give a multiplicity of products, each, however, possessing a potential aldehyde group with consequent increase in copper number. Complete hydrolysis must lead to glucose.

The long, molecular chain structure of cellulose, deduced from chemical evidence, is fully confirmed by X-ray analysis. When exposed to a monochromatic X-ray beam, a cellulose
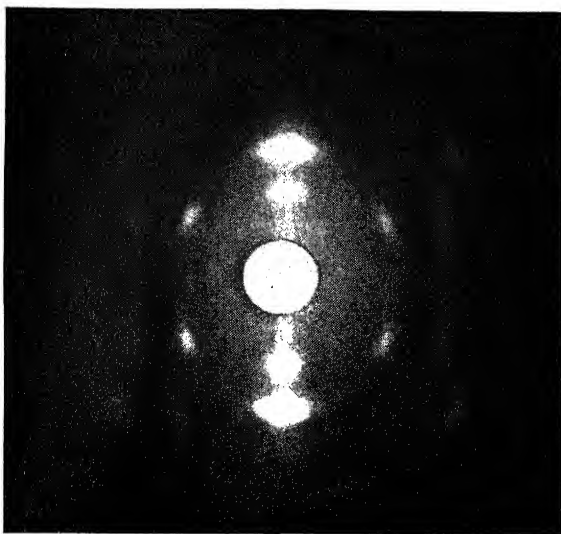
FIG. 3.   X-Ray Diagram of Cellulose.*

fiber shows a relatively sharp diffraction pattern (Fig. 3), demonstrating a high degree of molecular orientation (see p. 273). Furthermore, two glucose residues, arranged as in the above formula for cellulose, have dimensions exactly equalling the 10.3 Å spacing along the fiber axis. The insolubility of cellulose and its lack of thermoplasticity must be attributed to the high cohesion between the chains. This can be diminished to some extent by esterification, the larger the esterifying group the lower being the temperature at which thermoplasticity becomes

* Courtesy of Dr. B. E. Warren, Department of Physics, Massachusetts Institute of Technology.

apparent. Similarly, the tensile strength of ester fibers is lower the larger the substituent group.

Under many conditions cellulose suffers changes in properties involving loss in tensile strength of its fibers, increased chemical reactivity and solubility in reagents, decreased viscosity of its solutions, and the like. This is often true even where no visible change in the original fiber or formation of water-soluble products occurs. The change is known as degradation. It is most easily induced by the action of strong acids, which are active at elevated temperatures even in dilute solutions, the product in this case being called hydrocellulose. This points to hydrolysis, indicating that reduction in chain length is the effective cause of degradation.* Another important mechanism of degradation is oxidation in solutions of high pH. Cellulose is peculiarly sensitive to oxidizing agents in alkaline solutions. Apparently the chains are attacked at intermediate points, broken, and converted into carboxylic acid bodies, readily soluble in alkali, so-called oxycellulose. Protection against oxycellulose formation is vitally important in all industrial manipulation of cellulose.

Rubber.

The essential constituent in natural rubber is an unsaturated hydrocarbon of empirical formula C_5H_8. The rubber has high molecular weight, as indicated by negligible vapor pressure lowering of its solutions and as determined by the ultracentrifuge. Destructive distillation yields mainly a mixture of isoprene,

$$CH_2=C(CH_3)—CH=CH_2,$$
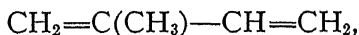
and certain of its derivatives. Rubber hydrocarbon contains one double bond for each C_5H_8 unit in the structure, as measured, e.g., by halogen addition. It was pointed out almost forty years ago that the chemical and physical characteristics of rubber can only be explained on the assumption that the hydrocarbon is an

---

* That the various changes by which degradation is recognized are due to a common cause is indicated by the fact that, starting with a given initial cellulosic material and degrading it by acids to various degrees, the tensile strength of comparable fibers corresponds closely to the viscosity of the same material dissolved at standard concentration in cuprammonium solution, solution in which under proper conditions has apparently little effect in breakdown of chain length.

.pen-chain polymer of isoprene of high molecular weight, in which one double bond of the isoprene has disappeared in the polymerization * (cf., styrene and isobutylene):

$$-CH_2-\underset{\underset{CH_3}{|}}{\overset{\overset{CH_3}{|}}{C}}=CH-CH_2-CH_2-C=CH-CH_2-CH_2-\overset{\overset{CH_3}{|}}{C}=CH-CH_2-.$$
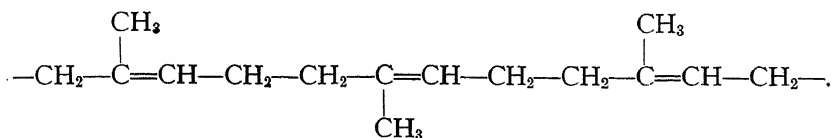
Polymerization of isoprene in the laboratory with various catalysts produces rubber-like substances, which, however, are not identical with natural rubber. Clearly, in this case, as in so many others, the conditions of polymerization in nature differ from those available in the laboratory.

Three-Dimensional Polymers. Polyesters.

So far, only simple bifunctional reactions have been considered, which produce either linear polymers or cyclic structures. When one (or both) of the reacting molecules possesses more than two reactive groups, the resulting products are much more complex and differ markedly from those discussed above. In illustration, compare and contrast the reactions with phthalic anhydride,

$$\overline{CO.C_6H_4.CO-O},$$
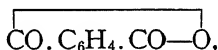
of ethylene glycol and glycerol. The former is similar to that between ethylene glycol and succinic acid, reaction going to 50% of completion at 190° C. in less than three minutes, but requiring about 2000-fold as long for 99% esterification. The flow point of the product, the temperature at which distortion under gravity begins to be rapid, rises only a few degrees for esterifications beyond 90%. At room temperature the product is a hard, glassy solid, which, however, softens reversibly on heating.

While substitution of glycerol for ethylene glycol gives a reaction similar in many respects, there is one outstanding difference. Up to about 80% esterification, flow point of the product rises slowly and progressively, but at about this point the mass sets to a gel structure, which increases in hardness with con-

* For a fuller discussion of the molecular structure of rubber, see pp. 404–410, 417–424.

tinued heating, and can never again be liquified by heat without decomposition. Furthermore, the final product is insoluble in solvents, and shows very little swelling in them. Increase in temperature greatly increases the rate of esterification, but the degree of esterification at the point of gelation is practically unaffected. While the molecular weight of the material just prior to gelation is sufficiently high to be difficult to measure accurately by cryoscopic methods, careful determinations indicate that it is relatively low, of the order of 2000 to 3000. After gelation, the insolubility of the resin makes molecular weight determination impossible.

Any linear polymer produced from a bifunctional molecule * is incapable of further molecular growth except by interaction of a functional group at the end of its molecular chain with a terminal group on some other molecule. While theoretically this can occur by interaction of two molecules, both of which are already large, it is clear that the probability of such a reaction is remote in comparison with that of the addition to the end of one long molecule of another short one.† In other words, the rate of increase of molecular weight with chemical reaction depends upon the degree to which reaction has already proceeded.

The mechanism of polymerization in a case such as glycerol phthalate may be extraordinarily different. Assume, for example, that a chain polymer of considerable length has already been synthesized by interaction of one glycerol molecule for each phthalic anhydride molecule. There must still remain along the chain residual hydroxyl groups of the glycerol radicals, any one of which is capable of interaction with a phthalic anhydride molecule. Even granting that, when an anhydride molecule reacts, it must react with two hydroxyls simultaneously, there is still abundant possibility for interaction of a given phthalic anhydride with glycerol hydroxyls, even though the anhydride molecule is not near the end of a polymer chain. Furthermore, interaction may combine together two chains, both of which are

---

* In this case cyclization due to the interaction of the functional groups clearly eliminates the possibility of further molecular growth.

† Even this rate should be low relative to the initial rate of reaction. The data in the case considered do, in fact, show that, computed in the conventional manner, the reaction is of approximately the fifth order, *i.e.*, rate falls extremely rapidly as reaction proceeds.

already long, thereby inducing an increase in molecular weight as the result of the esterification of a single anhydride molecule out of all proportion greater than can, on the average, be expected for interaction of the anhydride with a structure such as ethylene glycol.   Furthermore, assuming a mass already polymerized to a moderate degree, an amount of chemical interaction very small relative to the total which has previously occurred is capable of cross-linking the polymers already existing in the mass so that each is connected with its neighbor by one or more chemical valence bonds, leaving no chemically separate molecule in the structure.   In effect, this relatively small amount of chemical interaction can transform the whole particle into a single molecule, in the sense that one can pass from any point in the mass to any other point by traveling along primary valence bonds.

The synthetic resins may be divided into two general classes —the thermoplastic resins, which soften and liquefy with heat,* and the thermosetting resins, which under heat are transformed into non-fusible, insoluble solids.   In general, simple, bifunctional molecules polymerize to form the thermoplastic type of resins, whereas the thermosetting type results only from polymerization of simple molecules at least one of which contains a larger number of functional groups.

Striking confirmation of this interpretation of the cause of the fundamental difference between polymers of the thermoplastic and the thermosetting type is found in the behavior of styrene polymerizing in the presence of small quantities of p-divinyl benzene.   Polystyrenes of sufficiently high molecular weight are strong, tough, thermoplastic solids, swelling strongly before dispersion in suitable hydrocarbon solvents.   If, however, a little p-divinyl benzene (0.002 to 0.01%) is added to the styrene prior to polymerization, while there is practically no change in the initial stages of the reaction, a polymer is obtained which only swells and cannot be dispersed in solvents.   The larger the percentage of this material in the reacting mixture, the more marked is the effect.   The presence of the two double bonds in the divinyl benzene obviously makes it possible to develop long

* This is reversible if the softening point is below decomposition temperature.

branch chains and cross-linkages:

$$-CH-CH_2-CH-CH_2-CH-CH_2-CH-CH_2-CH-$$

$$-CH-CH_2-CH-CH_2-CH-CH_2-CH-CH_2-CH-$$

$$-CH-CH_2-CH-CH_2-CH-CH_2-CH-CH_2-CH-$$

$$-CH-CH_2-CH-CH_2-CH-CH_2-CH-CH_2-CH-$$
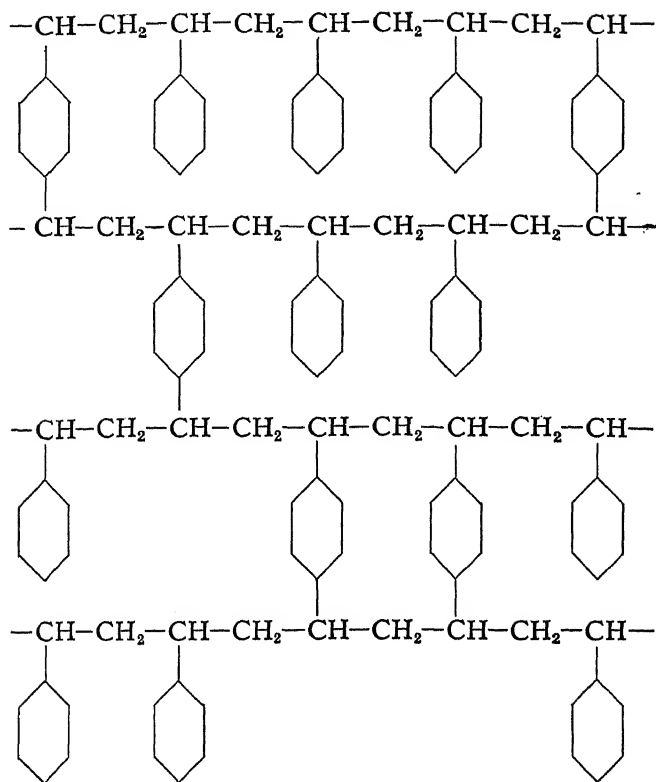
In view of the well-recognized tendency of organic materials of a given type to become more inert and decrease in chemical reactivity as molecular weight increases, it is not surprising to find that the synthesis of these so-called three-dimensional polymers offers one of the best means of preparing insoluble, highly resistant organic solids. For illustration of this industrially important group of solids the reader is referred to Chapters XX and XVIII.

An extraordinarily complicated type of condensation, difficult of general classification, is found in the proteins.* General treatment of these is beyond the scope of this text, but certain of them must be considered because of their practical importance,

* Current progress in this field is particularly rapid.

Proteins.

All living organisms contain nitrogenous substances, known as proteins, essential to vital processes. Many are of the form of typical amorphous solids, yielding colloidal solutions.

Proteins on hydrolysis with hot, concentrated hydrochloric acid yield predominantly α-amino acids,* of the general formula

$$R—CH(NH_2).COOH,$$

and amphoteric ring nitrogen compounds. In addition, some give substances † such as carbohydrates, lecithin, phosphorus derivatives, and the like.

Over twenty α-amino acids, of varying degrees of complexity, have been isolated by the hydrolysis of proteins. Some of the more important with their formulae are included in Table II.

Quantitative isolation and identification of all the amino-acids occurring in a single protein are difficult and perhaps never complete. Proteins differ widely in the amounts and character of the constituents (Table III).

Like cellulose, the proteins possess high molecular weights. This is evident not only from the failure of osmotic methods for determination but also from the figures in the above table.‡ In recent years the molecular weights of many proteins have been determined by Svedberg, using the ultracentrifuge (p. 107), yielding values varying from 34,500 for egg albumin to 5,000,000 for the respiratory protein of certain marine snails. It is a curious and as yet unexplained fact that the molecular weights appear to fall into two classes, those in the first ranging from 34,500 to 210,000 in multiples of 34,500, and those in the second with values in the millions.

When proteins are treated with nitrous acid, the amount of nitrogen liberated corresponds closely to one-half the amino nitrogen of the arginine and lysine present in the protein. $R.NH_2 + HNO_2 \rightarrow ROH + N_2 + H_2O$. Further indication of the presence of free amino and carboxyl groups is afforded

---

* So-called simple proteins yield only these.

† This section will not consider these last, conjugated or derived proteins.

‡ Thus, if the gelatin molecule be assumed to contain only one molecule of histidine, its molecular weight would be $100 \times 155/0.9 = 17,300$. Similarly, keratin and fibroin show minimum molecular weights, based on their histidine content, of 26,000 and 22,000, respectively.

TABLE II.   HYDROLYSIS PRODUCTS OF PROTEINS

A. Monobasic monoacidic amino acids.
   1. Glycine (aminoacetic acid) $CH_2NH_2.COOH$.
   2. Alanine (α-aminopropionic acid) $CH_3.CHNH_2.COOH$.
   3. Serine (α-amino β-hydroxypropionic acid) $CH_2OH.CHNH_2.COOH$.
   4. Phenylalanine (α-amino β-phenylpropionic acid) $C_6H_5.CH_2.CHNH_2.COOH$.
   5. Tyrosine (α-amino β-parahydroxyphenylpropionic acid)

$$p\text{-}HO.C_6H_4.CH_2.CHNH_2.COOH.$$

   6. Cystine [di- (β-thio-α-aminopropionic acid)]

$$CH_2\text{—}S\text{—}S\text{—}CH_2$$

$$CHNH_2 \qquad CHNH_2$$

$$COOH \qquad COOH$$
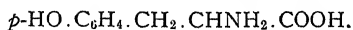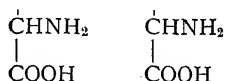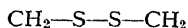
   7. Valine (α-amino isovaleric acid)$(CH_3)_2.CH.CHNH_2.COOH$.
   8. Leucine (α-amino isocaproic acid)$(CH_3)_2.CH.CH_2.CHNH_2.COOH$.
B. Dibasic monoacidic amino acids.
   9. Aspartic acid (aminosuccinic acid) $HOOC.CH_2.CHNH_2.COOH$.
   10. Glutamic acid (α-aminoglutaric acid) $HOOC.CH_2.CH_2.CHNH_2.COOH$.
C. Monobasic diacidic amino acids.
   11. Arginine (α-amino-δ-guanidine valeric acid)

$$NH_2$$

$$HN:C\text{—}NH.CH_2.CH_2.CH_2.CHNH_2.COOH$$
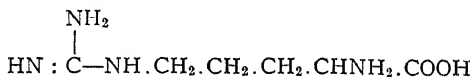
   12. Lysine (α-ϵ-diaminocaproic acid) $CH_2NH_2.CH_2.CH_2.CH_2.CHNH_2.COOH$
D. Heterocyclic Compounds.
   13. Histidine (α-amino β-imidazolepropionic acid)
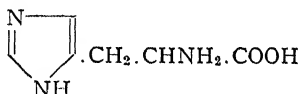
   14. Proline (α-pyrrolidine carboxylic acid)
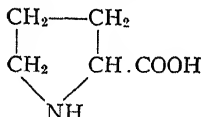
   15. Hydroxyproline (hydroxy-α-pyrrolidine carboxylic acid)
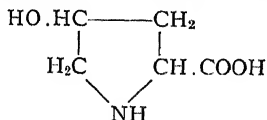
   16. Tryptophane (α-amino β-indole propionic acid)
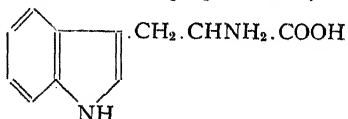
TABLE III. HYDROLYSIS PRODUCTS OF PROTEINS (WEIGHT AS PERCENTAGE OF ORIGINAL PROTEIN) *

Class	Constituents	Gelatin	Fibroin	Wool	Casein
A	Glycine	22	38	<1	—
	Alanine	6	23	4	1
	Serine	<1	2	3	<1
	Phenylalanine	1	2	—	3
	Tyrosine	—	11	5	5
	Cystine	—	—	13	—
	Valine	—	—	3	7
	Leucine	8	2	14	10
B	Aspartic acid	2	—	2	1
	Glutamic acid	6	—	13	16
C	Arginine	8	1	10	4
	Lysine	6	1	3	8
D	Histidine	1	<1	<1	4
	Proline	8	1	4	7
	Hydroxyproline	10	—	—	<1
	Tryptophane	—	—	2	2

by the amphoteric character of the proteins, which may be represented thus:

$$H_2N.R.COOH$$

acid   alkali

$$\overset{+}{H_3}N.R.COOH \qquad H_2N.R.CO\overset{-}{O}$$
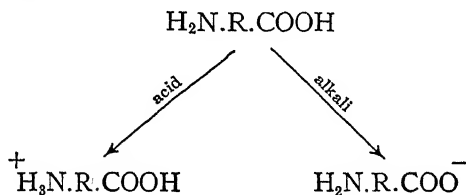
With gelatin, a typical protein, one equivalent of acid is found to combine with about 1100 gms. of gelatin, while an equivalent of base requires about 1700 gms. for neutralization. On the other hand, the average equivalent weight of the major amino acids in gelatin is approximately 100 as an acid and somewhat less as a base. This would indicate that over 90% of the acid and basic groups present in the original amino acids have disappeared when these substances are formed into proteins.†

* Data are approximate only. The variations in the results on what is presumably a single protein reported in the literature by different experimenters are often large. In important cases there is disagreement as to whether a certain amino acid is present in a given protein in major or minor amount. The variations are probably due in part to source of the protein investigated and in part to analytical difficulties.

† Actually, the acid and base combining values of the proteins can almost always be satisfactorily accounted for by the free amino or carboxyl groups of the diamino or dicarboxylic amino acids (classes B and C of Table II), as indicated by the behavior toward HNO_2.

The general mechanics of combination of amino acids to form proteins was first definitely established by Emil Fischer, who synthesized protein-like bodies having molecular weights of several thousand. He demonstrated that the α-amino group of one acid condenses with the carboxyl group of a second molecule, eliminating water and forming a peptide (—CO—NH—) linkage. Thus, two molecules of glycine may react as follows:

$$H_2N \cdot CH_2 \cdot COOH + H_2N \cdot CH_2 \cdot COOH \rightarrow$$
$$H_2N \cdot CH_2 \cdot CO \cdot NH \cdot CH_2 \cdot COOH + H_2O.$$

By virtue of the terminal carboxyl and amino groups, further condensation may occur, finally resulting in a long-chain polymer, a so-called polypeptide. The general formula may be written

—R—CO—NH—R′—CO—NH—R″—CO—NH—R‴—CO—NH—,

FIG. 4. X-ray Diagram of Silk Fibroin.*
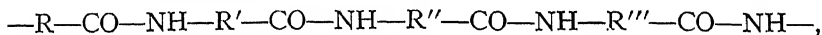
where R, R′, R″, R‴, etc. are amino acid residues.

This linear structure has been strikingly confirmed by the work of Astbury on the X-ray diffraction patterns of silk fibers. Purified silk fibers consist of the single protein, fibroin, over 65% of the known amino acids of which are glycine and alanine (Table III). Figure 4 shows the sharp X-ray diagram obtained. The predominant spacing is about 3.5 Å, corresponding exactly to the unit —CH—CO—NH—, spacing that has been found in many proteins which give X-ray patterns. Detailed analysis of the diffraction pattern indicates that fibroin can be written:

$$\overset{\displaystyle CH_3}{\underset{\text{Alanine residue}}{\underset{\displaystyle |}{-HN-\underset{\text{Glycine residue}}{CH_2}-CO-NH-CH-CO}}} \qquad CH_3$$

$$NH-CH_2-CO-NH-CH-CO-.$$
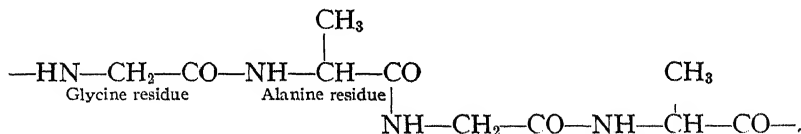
* Courtesy of Dr. G. L. Clark, Department of Chemistry, University of Illinois.

Although proteins apparently always contain polypeptide chains, there can be no doubt that some contain cross-linkages between the chains and are, in effect, three-dimensional polymers (p. 155).

Amorphous Form.

As the molecular weight of any specific type of primary valence polymer increases, the tendency to crystallize decreases. On the other hand, perfectly characterized crystals of high molecular weight substances are known. While a high degree of molecular complexity must retard the orderly arrangement which is essential in the crystal lattice, it seems improbable that this is a major cause of the resistance of high polymers to crystallization. On the other hand, the mechanism of synthesis of these polymers renders it highly improbable that the individual molecules are identical in structure. Indeed, it is frequently possible to fractionate a given polymer into products differing widely in molecular weight. While all the molecules of a given product may be identical in kind, their differences in molecular size would clearly tend to interfere with the degree of arrangement essential for the formation of a well-defined crystal lattice. Consequently, such materials resist crystallization sufficiently so that they can readily be supercooled, often remaining completely amorphous in the form of glassy, resinous solids. This fact is one of the most important reasons for their industrial utility (see pp. 277–278).

BIBLIOGRAPHY

BURK, R. E., THOMPSON, H. E., WEITH, A. J., and WILLIAMS, I., *Polymerization*, Reinhold (1937).

Collected Papers of W. H. Carothers on Polymerization, MARK, H., and WHITBY, G. S., ed., vol. I, Interscience Publishers, Inc. (1940).

FARADAY SOCIETY, *Symposium on Polymerization*, Sept., 1935.

HAWORTH, W. N., *J. Soc. Chem. Ind.* **54**, 839 (1935).

MARK H., *Physical Chemistry of High Polymeric Systems*, vol. II, Interscience Publishers, Inc. (1940).

MEYER, K. H., and MARK, H., *Der Aufbau der hochpolymeren organischen Naturstoffe*, Akad. verlag. m.b.h., Leipzig (1934).

STAUDINGER, H., *Die hochmolekularen organischen Verbindungen Kautschuk und Cellulose*, Springer, Berlin (1932).

Chapter VIII

Emulsoids

The Viscosity of Emulsoidal Solutions.

Solutions of many materials of high molecular weight (most of them amorphous solids) exhibit three flow characteristics which differentiate them sharply from ordinary liquids. In the first place, they show viscosities relative to that of the solvent which, even in dilute solutions, are very high; * in the second, the viscosity of a given solution may vary greatly with the rate of shear (so-called anomalous flow, see p. 167); and finally, under proper conditions even the dilute solutions often set to gels. Because the solutions exhibit properties in many ways analogous to those of emulsions of one liquid in another, they are called emulsoidal and the materials forming them, emulsoids.

It is obvious that the Einstein equation (p. 131) cannot explain high relative viscosities in dilute solutions. However, granting the general validity of the reasoning upon which the equation is based, it can still be used in solutions of this type if one assumes that, for some reason, the effective volume of the dispersed particles is much greater than their volume in the dry state. The Einstein equation may be rearranged:

$$2.5 \, V = \frac{\mu}{\mu_0} - 1 = \mu_r - 1. \tag{1}$$

Staudinger † has defined $\mu_r - 1$ as the specific viscosity.‡ It is clearly dependent on the concentration of the solution. Con-

---

* Thus, whereas solutions of 1% by volume of such substances as sugar, colloidal sulfur, or rubber latex have viscosities which are only 3 or 4% higher than that of the solvent, a similar solution of starch has a 50% higher viscosity, of agar over 500%, and of rubber over 1000%.

† STAUDINGER, H., *Die Hochmolekularen Organischen Verbindungen*, Springer, Berlin (1932).

‡ This terminology is unfortunate, since the quantity not only is dimensionless but is not even proportional to the viscosity.

sider a series of polymers of varying molecular weight but built up of identical chemical units [e.g., the polystyrenes $(C_6H_5 \cdot C_2H_3)_n$]. Neglecting differences in density between the polymers and granting that their dispersions conform to the Einstein equation, all such solutions should give identical viscosities at equal volumetric concentrations, independent of the molecular weight of the polymer. Staudinger has demonstrated that this is far from true.

The effective volumetric concentration of a given polymer should be at least approximately proportional to the concentration c expressed in gram-unit-groups per liter (i.e., for the polystyrenes this unit concentration is 104 gms. per liter). The Einstein equation would require that

$$\mu_r - 1 = 2.5 \, cv \qquad (2)$$

where v = the effective volume of one gram-unit-group. One would expect the term $2.5 \, v$ to be substantially constant for a given polymer series. Staudinger * has demonstrated that for polymers of the linear type this is not true, but that this term is proportional to the molecular weight of the polymer.† Hence, one should write the equation as

$$\mu_r - 1 = KMc \qquad (3)$$

where μ_r is the viscosity of the solution relative to that of the solvent, M the molecular weight of the chain polymer, and c the concentration in gram-unit-groups per liter. Even among the polymers of a given starting material (e.g., ω-hydroxydecanoic acid), K varies with M. Apparently, however, the variation is limited for molecular weights above 5000. Furthermore, K varies considerably with the material polymerized; e.g., $K = 0.0012$ for cellulose, 0.0011 for ω-hydroxydecanoic acids, and 0.000085 for paraffin hydrocarbons. However, if one divide K by the number of atoms, n, in direct line in each recurring unit of the chain (i.e., ignoring side chains), the result has been found roughly constant for very diverse types of polymers.

* Op. cit.

† In a given polymeric series, e.g., the polystyrenes, the molecular weight as thus determined by measurements of viscosity checked the results of cryoscopic and ebullioscopic determinations so long as the molecular weight was low enough for these last methods to be dependable.

Thus, since $n = 5$ for cellulose (p. 151) and $n = 1$ for paraffins, $K/n = 0.00024$ for the former and 0.000085 for the latter. For various types of hydrocarbons divergence is less. Thus, polystyrenes show $K/n = 0.00009$ and rubber, 0.000075. Type of solvent also affects the values. However, it is clear that these relationships make it possible to estimate the molecular weight of chain polymers by use of viscosity data, at least as a good approximation.*

Staudinger visualizes a chain polymer as a long rod which will tumble in the flowing liquid. In doing so chance of collision with similar rods is far greater than would correspond to its volume. This results in greater interference with flow, *i.e.*, in a greatly increased effective volume from the point of view of the Einstein equation. Assuming the effective volume of the rod proportional to the volume swept out by the rod rotating about an axis through its center perpendicular to its length, one can readily derive the Staudinger equation. There is no necessity to postulate rigidity of the rods, provided the effective rotational length is proportional to the number of atoms in the chain, and hence to the molecular weight.

The relation apparently applies generally to non-polar molecules of chain-like structure and known composition, but chain molecules containing ionizing groups, such as the sodium salts of the fatty acids, give anomalous values, apparently owing to molecular association into larger aggregates or micelles (see pp. 235–236).

It has never been possible to check the method independently in the case of materials of very high molecular weight. The best confirmation in this range has been by means of the ultracentrifuge (p. 107). While the method of Staudinger must be considered empirical, on account of simplicity of apparatus and ease of manipulation, it is a tool of extraordinary importance, particularly for laboratory control of polymerization.

* Applying the technique to the higher normal paraffins, fatty acids, and fatty acid mono-esters, Staudinger found that the specific viscosity of each series is an additive function of the number of organic residues in the molecules. Thus, for the fatty acids or their mono-esters,

$$\frac{\mu_r - 1}{} = km + a$$

where m is the number of methylene groups, CH_2, in the molecule and a and k are constants for a given series. Their obvious interpretation is that k is the contribution to the specific viscosity of an individual methylene radical and a, that of the terminal groups in the chain. For the hydrocarbons or similar materials, a is negligible, *i.e.*, the relation simplifies to (3).

As molecular weight of a given type of polymer increases, not only does specific viscosity increase but, in time, the solutions begin to exhibit anomalous flow similar to that of concentrated suspensions (p. 137). A plot of rate of shear against shearing stress is no longer a straight line through the origin, but, even at low Reynolds numbers where the flow is certainly laminar, the curve is concave upwards and may even approximate a straight line cutting the shearing force axis at a positive finite value. The relations are brought out in Fig. 1, based on data of Staudinger for polystyrene dissolved in tetralin. The dotted lines are for a polymer of low molecular weight at a concentration of about 30 weight per cent. The full curves are for a polymer of high molecular weight in a solution diluted to give approximately the same viscosity as the first.

FIG. 1. Rate of shear—shearing stress diagram for two polystyrene solutions (one full and the other dotted) of differing molecular weights. The upper curve of each pair is at 40° higher temperature than the lower.*

Three points should be emphasized: as is to be expected from the preceding discussion, the high molecular weight polymer gives approximately the same viscosity as the low at one-fifteenth the concentration; however, the high molecular weight material exhibits unmistakable anomalous flow, the viscosity of the solution decreasing progressively as rate of shear increases, in contrast to the normal behavior of the low molecular weight polymer at substantially identical viscosity; finally, whereas a 40° temperature fall gives a relative viscosity increase of 120% in the solution of the low molecular weight polymer, in that of the high molecular weight the increase is only 14%. In a stationary solution of a linear polymer, fully dispersed, the filament-like molecules are pre-

* Data of STAUDINGER, H., and HEUER, W., Ber. 62, 2933 (1929).

sumably in completely disordered arrangement. When flow starts, their mutual interference causes a high internal friction, producing a correspondingly high viscosity—predominantly by the mechanism pictured by Staudinger. However, streamline flow inevitably tends to comb out the polymer molecules, orienting them more or less parallel to each other and to the direction of flow. This is true, despite the fact that a certain fraction (even of oriented molecules) will be tumbled by having the front end caught in a slow moving lamina and the back end simultaneously in a neighboring, more rapidly moving one. In other words, the molecules will act much as logs going down a stream. However, this combing or orienting effect of flow is opposed by the effects of thermal agitation, tending to restore the original condition of disordered arrangement. The longer the molecule, the greater will be the combing effect of a given rate of slip of one lamina over its neighbor, relative to the disorganizing effect of thermal agitation. In the low molecular weight polymer of Fig. 1, at the rates of flow measured, no indication of chain orientation can be discovered. With the high molecular weight polymer, however, the viscosity falls progressively with increasing velocity gradient, indicating that at high velocities thermal agitation is insufficient to prevent effective orientation, causing corresponding decrease in viscosity.*

In solutions of the low molecular weight polymer, relative viscosity decreases rapidly as temperature rises but this relation is by no means general.† In the solutions of high molecular weight, the effect of temperature on relative viscosity is greatly suppressed, a phenomenon common among linear polymers. The orienting effect of flow on molecules of given length is presumably little affected by temperature. However, as molecular length gets high, the countereffect of thermal agitation becomes so small relative to the molecular interference with flow that the temperature influence is greatly reduced.

* For the calculation of molecular weight by Staudinger's method it is clear that one should employ data obtained under conditions where these orienting effects are negligible, since otherwise viscosity becomes a function of rate of shear, for which the Staudinger equation does not allow. This is done by operating at both concentrations and rates of shear sufficiently low that the curves of Fig. 1 are substantially straight lines.

† The cause in this case may be decrease of solvation (p. 170) of the ultimate molecules with temperature, but this is not certain.

In any solution in which dispersion is complete, one would anticipate a curve of the shape of those of Fig. 1, in which it would come to the origin even though with zero slope.   In solutions of the linear polymer type where the forces of thermal agitation are insufficient to disperse the polymer completely, there is, however, the possibility of an interlocking network or structure, the strength of which may be sufficient to prevent breakdown and flow under low loads.   Such a structure would have a finite yield value, below which flow would not occur, but above which the disruptive forces of shear superimposed on those of thermal agitation would be sufficient to induce it (see p. 224).

Some high molecular weight solids, particularly a number of the proteins, which might be expected to be linear polymers, approach Einstein's equation much more closely, and it seems likely that they possess spherical molecules.   X-ray diffraction patterns also support this view.

General Characteristics of Emulsoids.

The differences in properties of emulsoids * and suspensoids are brought out in Table I, which compares and contrasts a

TABLE I.   PROPERTIES OF SUSPENSOID AND EMULSOID SOLS

Suspensoids (*e.g., Gold Sol*)	*Emulsoids* (*e.g., Gelatin Sol*)
1. High molecular weight	High molecular weight
2. Low diffusion velocity	Low diffusion velocity
3. Can be purified by dialysis	Can be purified by dialysis
4. Cannot be prepared by direct solution	Prepared by direct solution
5. Stable only if charged or protected	Inherently stable
6. Stable only if dilute	Stable even if concentrated
7. Coagulation irreversible	Reversibly precipitated and dissolved
8. Optically non-homogeneous	Nearly optically homogeneous
9. Often colored	Usually colorless
10. Migrate in an electric field	May or may not migrate in an electric field
11. Not viscous	Generally relatively viscous even if dilute
12. Never form gels	Frequently form gels
13. Dry solid inert to solvent	Dry solid swells in solvent before dissolving

* Alternative names for emulsoids are lyophilic sols or hydrophilic (in the case of aqueous sols) sols (see p. 112).

typical suspensoid (gold sol) and a typical emulsoid (gelatin dissolved in water).

Since items 4 to 11, in the case of suspensoids, depend upon either the mechanism of stabilization or the structure of the particle, it is clear from the diametrically opposed behavior of emulsoids that the causes of their stability and structure must be very different.

Character of Emulsoidal Solutions.

When dry emulsoids are treated with solvent, unlike crystalline materials solution is preceded by a considerable degree of swelling, accompanied by imbibition of solvent. Indeed, at relatively low temperatures it is not uncommon for action to cease with the swelling stage, so-called limited swelling, the solid requiring rise in temperature before solution occurs (cf. p. 224).*

The initial swelling of the solids in the solvent is inescapable proof of positive affinity of the solvent for the solute. This affinity must be greater than that of the solute particles for each other, since otherwise they could not be separated by the solvent action. This suggests combination of solvent molecules at least with the surface of the solute particles, probably under the influence of the surface forces already discussed. This assumption immediately clarifies the stability characteristics of emulsoidal solutions. The solute particles, though large, will not agglomerate on contact, partly because the layers of solvent molecules held on their surfaces act as buffers and partly in view of the fact that even though contact occurs, it will be followed immediately by separation under the action of the forces causing swelling in the first place. Furthermore, the combination of solvent and solute just pictured will inevitably greatly increase the volume of the solute particle and, in the light of the Einstein relation, increase abnormally the viscosity of the solution. Finally, it would not be surprising to find highly solvated particles transformed to the extent of becoming themselves quasi-liquid and readily deformable, with a structure similar to that of a material such as rubber or gelatin, which has been greatly swollen in a solvent though not to the point of complete dissolu-

---

* At this stage solution can sometimes be induced by chemical action.

tion.   The interference of such suspended particles when under shear would, because of ready deformation, be expected to be less than that of rigid particles of equal volume.   In other words, a suspension of such particles might be expected to show decreasing interference to flow with increasing shear, *i.e.*, anomalous flow.   The facts point toward the assumption of solvation, *i.e.*, definite combination of the emulsoidal particles with solute molecules, as the underlying cause of many of the characteristic differences in behavior of emulsoidal solutions on the one hand and suspensoids on the other.

Emulsoid sols are pictured, therefore, as made up of large molecules or particles, heavily solvated, so that their volume is much greater than that in the dry state.   Frequently, but not necessarily, the particles possess an electric charge.   However, while further study will confirm the conclusion that solvation is a primary factor in the formation and stability of emulsoids, the phenomena are complicated in the extreme and other factors play an important part.

Precipitation and Flocculation of Emulsoids.

If the stability of emulsoids be due largely to the fact that the individual particle is highly solvated, then it is obvious that to precipitate the emulsoid from solution the solvent must be removed.   This may be done directly by evaporation, which usually regenerates the substance in its original amorphous condition, and, if solvent be again added, the sol reforms.   This is the usual technique when it is desired to use the emulsoidal solution as an adhesive or to apply it as a thin film over a surface, as in the case of varnishes and lacquers (Chapter XIV).

Granting solvation is an effect of the secondary valence attractions or van der Waals forces between the material and its solvent, it should follow the mass action law at least approximately, *i.e.*, the solvation process should tend to reverse on reducing the concentration of solvent, as by adding a liquid which is miscible with the solvent in all proportions but is a non-solvent for the amorphous material in question.   Thus, considering rubber in benzene as a typical example, if an appreciable amount of acetone, in which rubber is practically insoluble, be added, the

solvation of the rubber should be greatly reduced. In practice, the rubber is precipitated from solution by approximately 30% of acetone. The precipitation of hydrated emulsoids, such as gelatin, from aqueous solution by ethyl alcohol is another example. While the addition of sufficient liquid, non-solvent for the emulsoid but miscible with the solvent, will precipitate the solute, it would be surprising if its effect did not show up in modified properties of the solution before the precipitation point is reached. This is actually the case, as shown by the influence on the viscosity of the sol. Thus, while addition of a small amount of acetone to a rubber cement would be expected to cause a decrease in viscosity because of the dilution, it is found that the actual drop in viscosity is greater than would be anticipated on this ground. The desolvation, which addition of the non-solvent induces, should result in reduction of size of the emulsoid particles and this in turn in reduction of viscosity. However, the desolvation must be relatively complete before precipitation results. This effect of a miscible non-solvent on the viscosity of an emulsoid is quite general.

In the case of addition of electrolytes to emulsoid sols in polar solvents, similar phenomena are to be anticipated, but one must be on the lookout for superimposed ionic effects. Electrolytes are much more effective, per mol added, than non-electrolytes, such as acetone or alcohol, although, in contrast to suspensoids, hydrophilic sols possess a relative immunity toward them. While in most cases only a few millimols of electrolyte per liter are required to precipitate a suspensoid sol (p. 123),

TABLE II. MINIMAL MOLAL CONCENTRATIONS OF SALTS REQUIRED TO PRECIPITATE A 0.8% GELATIN SOLUTION *

pH of gelatin solution	$(NH_4)_2SO_4$	Na_2SO_4	$MgSO_4$	KCl	$MgCl_2$
4.7 (isoelectric point)	15/16 M	6/8 M	10/8 M	>3 M	>3 M
3.8	13/16	5/8	7/8	3	>3
6.4 to 7.0	16/16	7/8	9/8	>3	>3

much more electrolyte is necessary to flocculate hydrophilic sols (Table II). In many instances the solubility of the salt in question sets the limit. Thus, many protein sols are precipitated

* LOEB, J., *Proteins and the Theory of Colloidal Behavior*, McGraw-Hill (1922).

by the use of half-saturated magnesium sulfate and are almost entirely unaffected by small traces, which would be sufficient to produce complete coagulation of a suspensoid. In the latter case only enough electrolyte is required to furnish the ions needed to bring about effective neutralization of the particle charge. In the case of emulsoids, however, sufficient electrolyte is required to remove a large amount of water from the emulsoid particle, *i.e.*, the amount of electrolyte must be sufficient to compete successfully with the amorphous material for the solvent. Mass action effect requires a large amount to desolvate the dissolved particle sufficiently to precipitate it.

The addition of neutral electrolyte to an emulsoid often produces a large reduction in its viscosity (the electroviscous effect of Kruyt) similar to the effect of non-solvents. This reduction is undoubtedly due to a decrease in the total effective volume of the disperse phase, but whether this is caused by mere shrinkage of the individual particles or by their aggregation into large clumps is uncertain. Probably both effects are encountered and it is difficult to distinguish their relative importance. The change in character of the particles is confirmed by the fact that the addition of precipitating agents of this sort usually produces an opalescent condition in the sol, greatly increasing the brilliance of its Tyndall cone effect, whether observed directly or in the ultramicroscope.

That an emulsoid sol can be entirely stable in absence of an electric charge seems obvious, since it is difficult to picture polystyrene or pure linear hydrocarbon polymers of any kind dispersed in an organic liquid carrying a charge. Furthermore, a hydrophilic sol, *e.g.*, gelatin in water, can be entirely stable when it travels neither to anode nor cathode in an electric field, showing that it possesses no net electric charge. Nevertheless, in some emulsoids charge may be an important contributory cause of stability. Thus, small additions of $MgSO_4$ or $MgCl_2$ to an aqueous agar sol produce sharp drops in viscosity, identical for the two salts for equal molal concentration, probably due to the neutralization of the charge on the agar particles by the magnesium ion. * Von Smoluchowski has pointed out that Ein-

* Kruyt, H. R., *Colloids*, Chapman and Hall, Ltd., London (1930), p. 197.

stein's equation takes no cognizance of the charge on the particles, which would tend to increase their effective volume with the result that the viscosity of charged suspensions should be considerably higher.

At high concentrations of neutral electrolytes ionic hydration is a significant factor. It is well known that the polar electrolytic ions are highly hydrated in solution. Furthermore, the ions can be arranged in the order of their hydration, as in Table III, even though the actual amount of water associated with any of them cannot be determined with certainty. Such a list, called a "lyotropic or Hofmeister series," is of importance in connection with hydrated emulsoids. In general, the ions of

TABLE III. HYDRATION OF IONS ARRANGED IN ORDER OF
DECREASING HYDRATION

Cations	Li,	Na,	K,	Rb,	Cs	and	Mg,	Ca,	Sr,	Ba
Anions	SO₄,	Cl,	Br,	NO₃,	I,	CNS				

higher hydration are most effective in the "salting out" of emulsoids; the solubility of the salt is of secondary importance. Thus, in the present case, although magnesium chloride is more soluble than magnesium sulfate, the component ions of the latter are more highly hydrated. Consequently, the sulfate causes more dehydration of the agar; the viscosity breaks more rapidly and the sol finally flocculates at low concentrations. Magnesium chloride, on the other hand, does not dehydrate the sol so effectively and the sol remains stable up to a concentration of 4 molal. The action of the ions is not only due to their hydration but probably depends also upon the way that they affect the known association of the water molecules with each other, the ions probably displacing the equilibrium in favor of simple H_2O molecules.

This theory of the stabilization of emulsoids is confirmed by the behavior of the sols when treated with amounts of non-electrolytic dehydrating agents insufficient to precipitate them, the sol being "sensitized" towards electrolytes. Thus, whereas agar-agar in aqueous solution is a characteristic emulsoidal sol, such a solution diluted with five volumes of alcohol is trans-

The anions should read: SO$_4$, Cl, Br, NO$_3$, I, CNS

formed to one substantially suspensoidal in type. While this alcoholic solution is still stable, the addition of a little over a millimol per liter of KCl will induce flocculation; furthermore, only 0.05 millimol per liter of barium chloride is required to flocculate, and only 0.03 millimol of $Co(NH_3)_3Cl_3$ is necessary. In other words, the sol containing alcohol is not only very sensitive to electrolytes but the valency of the precipitating ion is of great importance. Moreover, the initial aqueous sol can be sensitized by addition of the electrolyte, and flocculated by means of alcohol, *i.e.*, the order of addition of the modifying materials is of minor importance.

As has been emphasized by Kruyt, many emulsoidal sols can be rendered less stable, often to the point of flocculation, either by desolvation (by addition of a liquid miscible with the solvent but a non-solvent for the emulsoid), or by modification of charge on the colloid by addition of electrolyte, or by combination of both, the relative effects being largely controllable.

Protective Action.

Classification of colloids into suspensoids and emulsoids is not rigid. Many sols are known which exhibit some properties of both types and form a connecting link between the two classes. Furthermore, in many cases suspensoids can be converted into sols having certain properties characteristic of emulsoids by the addition of small quantities of emulsoidal material.

If gelatin be added to a gold or silver sol, which has been prepared in such a way that it is alkaline in reaction, it is found that a greater quantity of electrolyte is required to precipitate it than before. The physical appearance of the sol is unchanged, but it has become more stable. In some cases metallic sols can, after addition of a suitable emulsoid, be evaporated to dryness and redispersed in water without difficulty. Because of this decreased sensitivity the suspensoid sol is said to be "protected" and the phenomenon is called "protective action." Many other emulsoidal colloids besides gelatin behave in exactly the same way. Zsigmondy, who first systematically studied protective action, devised an empirical method for comparing the protective power of different emulsoids. He defined the "gold number"

of a protective colloid as the number of milligrams of it just sufficient to prevent a change in color from red to blue of 10 c.cm. of gold sol, upon the addition of 1 c.cm. of a 10% sodium chloride solution. In practice the gold number is found by a process of trial and error under carefully standardized conditions.* Values obtained by Zsigmondy and Gortner for a number of protective colloids are shown in Table IV, the qualitative trend being the same, although differing somewhat numerically. It must be emphasized that the method is comparative, serving only as a rough measure of protective action. A number of factors influence the results, e.g., the degree of dispersion of both the gold sol and the protective colloid, the pH of the solution, the dilu-

TABLE IV.   GOLD NUMBERS OF PROTECTIVE COLLOIDS †

Colloid	Gold Number	
	Zsigmondy	Gortner
Gelatin	0.005–0.01	0.005–0.0125
Isinglass	0.01–0.02	———
Casein	0.01	———
Egg albumin	———	0.08–0.10
Protalbinic acid	0.03–0.08 (Na salt)	0.15–0.20
Lysalbinic acid	0.02–0.06 (Na salt)	0.10–0.125
Gum arabic	0.15–0.5	0.10–0.125
Gum tragacanth	About 2	———
Dextrins	6–20	———
Dextrin (British Gum)	———	125–150
Soluble starch	———	10–15
Potato starch	About 25	———
Sodium oleate	0.4–1	2–4

tion, etc. The magnitude of the effect of dispersion of the colloid is shown by the fact that it may require ten times as much gelatin added in the form of a 1% solution as is necessary if the original gelatin concentration be 0.001%. This points to

* A definite amount of the protective colloid in solution is added to the gold sol, the mixture well shaken for three minutes, and the sodium chloride run in while stirring. A color change indicates that sufficient protective colloid to prevent incipient coagulation (cf. p. 117) has not been added and a larger amount must be used in the next experiment.

Zsigmondy used a gold sol prepared by the reduction of chlorauric acid with formaldehyde in alkaline solution. The sol should have a clear, bright red color by transmitted light, with no trace of violet or blue. After addition of protective colloid some time should be allowed to elapse before the electrolyte is added, protective action being not quite instantaneous. In general, three minutes is sufficient.

† ZSIGMONDY, R., Z. analyt. Chem. 40, 697 (1901).
   GORTNER, R. A., J. Am. Chem. Soc. 42, 595 (1920).

a high degree of aggregation of gelatin molecules, even in 1% solution.

Although a gold sol is often chosen for comparison of the protective powers of emulsoids, obviously this may vary greatly toward different suspensoids. Thus, while gelatin has approximately 10,000-fold the protective action of saponin on a gold sol, on one of As_2S_3 its protective effect is only 1/7 as great.

Investigation of the behavior of protected sols in an electric field throws light on the mechanism of stabilization. Particles in different sols to which excess gelatin has been added show the same velocity under similar conditions. Thus, the particles in a negatively charged manganese dioxide sol and those in a positively charged ferric oxide sol, to both of which an excess of
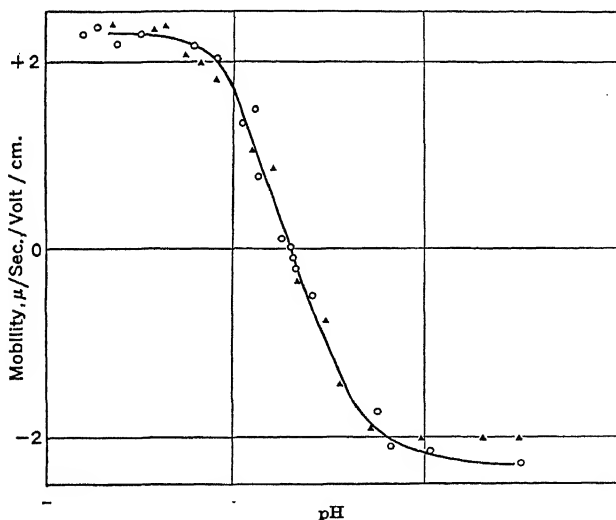
FIG. 2.   Cataphoresis of Protected MnO_2 and Fe_2O_3 Sols.*
O—Fe_2O_3 Sol, △—MnO_2 Sol

gelatin has been added, move with the same cataphoretic velocity at the same pH, as shown by Fig. 2 (in this connection, see pp. 206–207).

Protected sols behave toward electrolytes in a manner similar to that of the protecting colloid itself.   The Hardy-Schulze rule

(see p. 125) no longer holds, the valency of the precipitating ion
having comparatively little effect on the quantity of electrolyte
required for coagulation.   This is well shown by data of Loeb on
collodion sols.   Fine suspensions of collodion in water may be
prepared by adding water to an acetone solution of collodion,
removing the acetone by distillation under reduced pressure,
centrifuging, and rediluting with water.   The sols can be pro-
tected by allowing the collodion particles to stand in a suspension
of the emulsoid overnight, centrifuging, and rediluting in water
of the desired pH.   Table V shows the minimal concentrations
of various salts required to flocculate a pure collodion suspension
within 18 hours at various hydrogen-ion concentrations.   With-
out doubt, the collodion forms a negative hydrophobic sol.   The

TABLE V.   COAGULATION OF COLLODION SOLS *

pH	$NaCl$	Na_2SO_4	$CaCl_2$	$LaCl_3$
3.0	M/2	M/4	M/32	M/2,048
5.8	M/2	M/4	M/32	M/2,048
11.0	M/2	M/4	M/32	——

figures may be compared with those for a gelatin-protected
collodion suspension, the molar concentrations required to pre-
cipitate the suspension on standing overnight being given in
Table VI.   Much larger quantities of electrolyte are required
and evidence of a valence effect is apparently lacking.

TABLE VI.   COAGULATION OF GELATIN-PROTECTED
COLLODION SOLS *

pH	$NaCl$	Na_2SO_4	$CaCl_2$	$LaCl_3$
3.0	2 M	1 M	>2 M	>1 M
4.0	>2	1	>2	>1
4.7	>2	>1/2	>2	>1
5.8	>2	1	>2	>1
11.0	>2	1	2	

In Table VII, for purposes of comparison, are shown the
minimal concentrations of various salts required to salt out or
precipitate a 1% solution of gelatin, adjusted to the required
pH by the addition of acid or alkali.   The necessary concentra-

* LOEB, J., *J. Gen. Physiol.* 5, 479 (1922–3).

tions are almost identical with those required for the protected collodion suspensions.

The conclusion seems inescapable from these and similar data that protective action is due to adsorption of the protective

TABLE VII.   COAGULATION OF A 1% GELATIN SOL *

pH	$NaCl$	Na_2SO_4	$CaCl_2$	$LaCl_3$
3.0	2 M	1/2 M	>2 M	>7/8 M
4.0	>2	1/2 or above	>2	>7/8
4.7	>2	1/2 or above	>2	>7/8
5.8	>2	7/8	>2	>7/8
11.0	>2	7/8	>2	

colloid by the hydrophobic particles, so that an envelope or skin of hydrophilic material is formed around them.   Thus, the suspensoid particles become coated with a thin film having a high affinity for the solvent and assume to a corresponding degree the properties of the emulsoid.   The specificity of the protective power, a given emulsoid protecting only certain suspensoids, also points to an adsorption mechanism.   The adsorption of gelatin by gold can be shown directly by an experiment due to Zsigmondy, in which a thin sheet of gold is immersed in a gelatin solution and well washed with water.   Even after washing with hot water the sheet does not amalgamate with mercury.   In other words, the stability of the particles in the protected suspensoidal sol depends not upon their electric charge, but upon the affinity of the enveloping protective colloid for the solvent.

Although attention has been focused upon protective action in aqueous solution, undoubtedly the same phenomenon may occur in non-aqueous dispersion mediums.   Thus, white lead suspended in oil is probably protected by the lead salts of fatty acids present in the oils.

Sensitization.

In acid solution small quantities of gelatin may cause flocculation, still larger quantities causing protection of the gold particles.   Thus, the protective action of many emulsoids varies greatly with the pH of the solution, illustrated by experiments of

* LOEB, J., *J. Gen. Physiol.* 5, 479 (1922-3).

Reinders and Bendien, in which the influence of gelatin upon the stability of a gold sol was examined at a number of different hydrogen-ion concentrations. In Table VIII the appearance of the gold sol eighteen hours after mixing with the gelatin solution is given; the cataphoretic velocities (p. 188) of the particles were also determined and are shown. In acid solutions the positively charged protein first induced coagulation, paralleling that of arsenious sulfide and ferric hydroxide sols (p. 124), followed by sign reversal of the charge on the gold particles when further amounts of gelatin were added. In alkaline gold sols gelatin has the same sign as the gold particles (see p. 119), and hence shows normal protective action. In other words, the precipitating power of emulsoids is apparently a particular case of mutual coagulation. When smaller amounts of emulsoid are added, the suspensoid becomes much more sensitive to precipitation, *e.g.*, by electrolytes, and is said to be "sensitized."

TABLE VIII.   THE EFFECT OF GELATIN ON COLOR AND
CATAPHORETIC VELOCITY * OF GOLD SOLS †

pH	Gelatin Concentration (mgm./liter)					
	0	0.00005	0.0005	0.005	0.05	0.15
11	Red −5.0	Red −4.7	Red −4.2	Red −3.3	Red −2.2	Red −2.2
7.7	Red −4.2	Red −3.7	Red −3.0	Red −2.3	Red −1.7	Red −1.7
5.2	Red −3.9	Red violet −3.4	Violet −2.1	Red −1.5	Red −1.1	Red −1.0
5.0	Red −3.8	Red violet −3.3	Violet then coagulated −1.6	Red violet −0.6	Red −0.6	Red −0.6
4.7	Red	Red violet	Blue pptd.	Violet	Red	Red
4.2	Red −3.6	Violet −3.0	Blue pptd. −0.7	Violet +0.7	Red +0.9	Red +0.9
3.0	Red violet −3.4	Blue −2.9	Blue pptd. 0	Violet +2.1	Red +2.3	Red +2.2

Zsigmondy ‡ has compared the precipitating or sensitizing powers of various colloidal substances in exactly the same manner as their protective powers. The number of milligrams of col-

* Cataphoretic velocity of gold particles in μ/sec. for 1 volt/cm.
† REINDERS, W., and BENDIEN, W. M., *Rec. trav. chim.* **47**, 977 (1928).
‡ ZSIGMONDY, R., *Z. anorg. Chemie* **96**, 265 (1916).

loidal substance necessary to change the color of 10 c.cm. of a gold sol, to 100 c.cm. of which has been added 3.5 c.cm. of 0.1 N HCl, is called by him the U-number * of the substance.

Borderline Sols.

A large class of colloidal solutions exists, the stability of which is apparently about equally dependent upon solvation and electric charge. In this type, the behavior of the sol will be much more complex than in the case of the previously considered true emulsoids or suspensoids, solvation and charge masking their mutual effects. Compounds exhibiting this mixed behavior in pronounced degree are the oxides and hydroxides of the weakly acidic or amphoteric elements, e.g., silicon, tin, iron, chromium, and aluminum, certain types of clays, and specific emulsoids. The oxide sols can be prepared by methods which are essentially similar to those employed in the preparation of suspensoids (cf. p. 114 et seq.). Stannic acid, typical of this class, is preferably liberated by means of hydrochloric acid from dilute, aqueous stannate solution and purified by dialysis, or by pouring a dilute solution of stannic chloride into dilute ammonia. Sols prepared by acidification are usually positively charged, those prepared with alkali, negatively. Sols of the more acidic oxides, e.g., SiO_2, are best prepared by acidification of their alkali salts, followed by dialysis, while the more basic iron, chromium, and aluminum sols are usually prepared by hydrolysis of their chlorides.

If purification by dialysis be carried too far, the sol flocculates, since nearly all the stabilizing ion is removed. After flocculation, however, only a very small amount of electrolyte need be added to the precipitate in order to regenerate the sol or, as termed by Bancroft, to "peptize" it. As little as 0.0066 millimol of potassium hydroxide per gram of stannic oxide will bring the sol back into dispersion, in contrast to the difficulty of redispersing suspensoids. This action of electrolytes is sometimes reversible; thus, stannic acid sol can be precipitated by electrolytes and redispersed by their removal as often as desired. Hydration of the precipitated particles ap-

* From the German "Umschlagezahl"—inversion number.

parently is sufficient to keep them from permanently agglomerating into larger aggregates, which cannot be reseparated. In other words, the large particles formed when the sol precipitates are easily redispersed by removal of electrolyte, since the coalescence of the primary particles is not complete. While the action of electrolytes upon these sols is qualitatively similar to their action upon suspensoids, larger quantities are required and the effect of the valence of the discharging ion is less pronounced. To precipitate a stannic oxide sol requires about twice the concentration of KCl needed to flocculate a gold sol. For $AlCl_3$, however, this ratio is about 400, although $AlCl_3$ is still 40 times as effective on stannic oxide as KCl.

The relative viscosity of sols of this class is higher than for suspensoid sols, indicating a larger effective volume of the colloidal particles and hence probable hydration. If this be true, any factor lowering hydration should reduce the relative viscosity of the sol and, hence, the amount of electrolyte required to flocculate it. Table IX, in which the effect of adding alcohol to a ferric hydroxide sol is given, shows this to be the case. Similar behavior was referred to in the case of agar sols (p. 174).

TABLE IX.   ACTION OF ALCOHOL UPON A FERRIC
HYDROXIDE SOL [a]*

Conc. Alc. Wt. %	Viscosities of		Rel. Vis. of Sol	Millimols of NaCl to Cause Complete Precipitation, per Liter
	Sol+Alcohol	Water+Alcohol		
0.0	1.018	1.000	1.018	22
10.2	1.478	1.459	1.013	16
20.4	2.050	2.029	1.010	12
43.1	2.652	2.635	1.006	8
54.9	2.547	2.530	1.007	7

[a] Fe_2O_3 = 2.0 gm./liter.

Investigation of the viscosity of the oxide sols discloses a phenomenon not shown by ordinary emulsoids and suspensoids, namely, ageing, viscosity serving as a convenient index. Ageing also causes an increased tendency to flocculation by electrolytes. Generally, this change in properties takes place relatively rapidly

soon after preparation, slowing up as time goes on. Under ordinary conditions the ageing process may take many months to make itself felt but the rate increases rapidly with temperature. The effect of temperature on a ferric hydroxide sol is shown in Table X. The cause of decreasing stability is probably

TABLE X. THE EFFECT OF TEMPERATURE ON A FERRIC HYDROXIDE SOL[a]*

Temp. Heated to, °C.	Rel. Visc.	Floc. Value
100	1.009	15
140	1.007	12
160	1.006	8
170	1.002	2
175	1.001	1
180	1.000[b]	0[b]

[a] $Fe_2O_3 = 1.0$ gm./liter.
[b] This sample completely precipitated by heating.

particle dehydration. When a silicic acid sol is dialyzed immediately after preparation, about 20% of the total silicic acid present passes through the membrane. On keeping the sol for some time and then dialyzing, practically no silica diffuses through. At the same time the molecular weight measured by the cryoscopic method shows a progressive rise (Table XI). In some sols changes of color also occur during ageing.

The oxide sols resemble emulsoids in forming definite, well-characterized gels, although these differ in being irreversible, i.e., once formed or set they do not, in general, exhibit swelling or redisperse in the solvent (pp. 225–228).

It seems clear that these sols must be regarded as intermediate in properties between the emulsoids and suspensoids, being stabilized by both electric charge and solvation, but with the phenomena complicated by the time changes which cause ageing. The silicate ion in a solution of sodium silicate is probably solvated with a large amount of water, as is the case for most ions. On acidification, the silicic acid is, in all probability, liberated as individual molecules possessing the water of hydration formerly held by the silicate ions. It is well known,

* Ayres, G. H., and Sorum, C. H., *J. Phys. Chem.* **34**, 2826 (1930) (The Williams and Wilkins Co.).

however, that weak acids tend to form complexes * in which oxygen-bridging occurs. Thus, it appears probable that the individual silicic acid molecules condense by elimination of water between two silicic acid residues in somewhat the following manner.
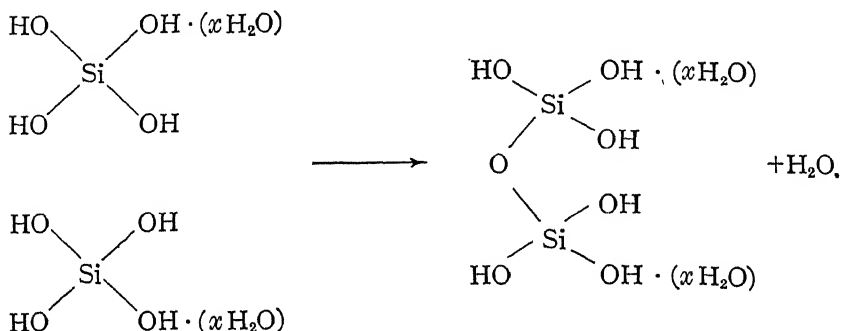
x H$_2$O representing the water of hydration. Condensation can then proceed further by a similar mechanism. The water of hydration of the polymer formed, as distinguished from the water eliminated by chemical condensation, is apparently not lost immediately but slowly diminishes. Finally, continued condensation produces large molecules of colloidal size, which may be represented as
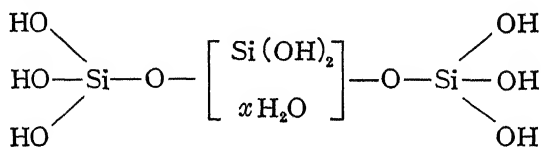
The bracketed portion represents the condensed, unionized silicic acid with its water of hydration very loosely held, while the radicals outside the bracket represent the end groups of the chain. The formation of other sols of this type, such as stannic acid, ferric hydroxide, aluminum hydroxide, and the like, may be similarly represented. The hydroxyl groups on each condensed chain can react with acid or alkali, as the case may be, causing the charge on the particle. The water of solvation,

* Chloric and nitric acids form no such salts. Sulfuric acid forms pyrosulfates, *e.g.*, K$_2$S$_2$O$_7$, which are, however, known only in the solid state, decomposing in water. Phosphoric and boric acids form many complexes of this type which are stable even in solution.

however, as represented by the water inside the bracket, tends progressively to be given off, thus decreasing the hydration of the particle. This is another important distinction between the oxide sols and the true emulsoids. In the case of the latter the water of hydration is firmly held and there is little tendency to progressive dehydration. On the contrary, the oxide sols tend to lose their water of hydration with time, becoming more and more suspensoidal in character. Thus, ageing consists of a gradual transition from a low molecular weight sol, predominantly emulsoidal in type, to a high molecular weight sol of suspensoidal character. Even more significant is the fact that there remain distributed along the chains reactive hydroxyl groups from which still more water can be eliminated, either by interaction of neighboring hydroxyls on a given chain or of hydroxyls on adjacent chains. This latter effect produces cross-bridging between chains (see p. 155), resulting in three-dimensional polymerization by stable, primary valence bonds.

TABLE XI.   CHANGE IN VISCOSITY WITH TIME OF A
SILICIC ACID SOL *

Age of Sol		Viscosity	Mol. Wt.
Hrs.	Mins.		
0	0	2.0	159
110	0	2.0	1500
	10	2.45	
	20	3.2	
	30	4.7	
	40	8.5	
	45	14.5	
	47	21.0	
	50	51.0	
	51	115.0	
	52	342.0	
	53	605.0	
	54	1144.0	
	55	1683.0	
		Gelled	

The parallelism in behavior between a sol such as silicic acid and the setting of polymers known to possess this type of struc-

* TREADWELL, W. D., *Trans. Faraday Soc.* **31**, 297 (1935).

ture (p. 473) is brought out in Table XI, for salt-free 0.5 M silicic acid sol of pH 3.2. This sol showed no detectable increase in viscosity for 110 hours, during which period, however, its molecular weight, cryoscopically determined, rose almost ten-fold. It then started to set rapidly, gelling in about an hour.

BIBLIOGRAPHY

FREUNDLICH, H., *Kapillarchemie*, Ed. 4, Akad. verlag. m.b.h., Leipzig (1930).
HOUWINK, R., *Elasticity, Plasticity and Structure of Matter*, University Press, Cambridge, England (1937).
SVEDBERG, T., *Colloid Chemistry*, Chem. Cat. Co., Ed. 2 (1928).
THOMAS, A. W., *Colloid Chemistry*, McGraw-Hill (1934).

Chapter IX

The Electrochemical Behavior of Colloids

Reference has already been made to the movement of colloidally dispersed particles in an electric field. The phenomenon is known as cataphoresis or electrophoresis and is common to gases as well as liquids.* It indicates the presence of forces between the phases, which can induce relative motion under widely differing conditions, producing different kinds of electrokinetic effects. That this is the case is shown by a simple but effective experiment, first performed by Reuss in 1808 (Fig. 1). Two open glass tubes are embedded in a mass of wet clay, a layer of coarse sand being placed at the bottom of each tube to prevent mechanical disturbance of the clay particles. Water is poured into the tubes, electrodes inserted, and a potential difference applied. The water level rises in the cathode and falls in the anode compartment, clay particles simultaneously rising towards the positive pole. This is a clear example of movement both of liquid relative to the solid (electroendosmosis) and of solid relative to the liquid (electrophoresis).
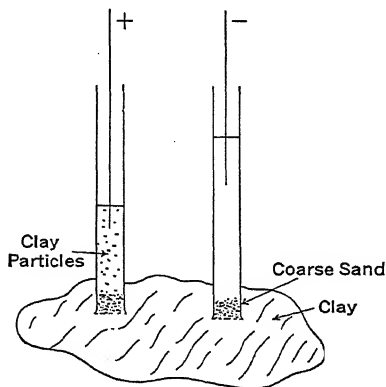
FIG. 1. Reuss' Experiment.

Conversely, it is found experimentally that relative motion of two phases under suitable conditions often induces a potential difference. Thus, there are four electrokinetic phenomena possible, listed in Table I.

In order to account for these effects, Helmholtz introduced the concept of an electrical double layer at the interface between

* In gases it is the basis of the Cottrell precipitator for dusts.

187

the two phases, somewhat as shown in Fig. 2(a) and (b).   In these, the positive charges are pictured as rigidly attached to the

TABLE I.   ELECTROKINETIC PHENOMENA

Solid Phase in Form of	Phenomenon	Cause	Name
Wall of tube	Liquid moves through tube	Applied E.M.F.	Electroendosmosis
Particles	Particles move through liquid	Applied E.M.F.	Electrophoresis
Wall of tube	E.M.F. set up	Liquid flowing past wall	Streaming potential
Particles	E.M.F. set up	Particle moving through liquid, e.g., under gravity	Migration potential (Dörn effect)

solid, the negative charges being free to move with the liquid.* Application of a potential difference to the system in Fig. 2(a)

(a)

(b)
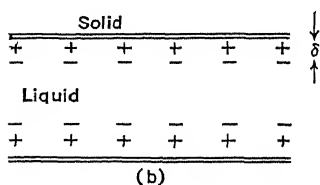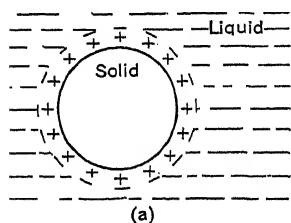
FIG. 2.   The Electric Double Layer.

would cause the particle to move toward the negative electrode; on the other hand, if the solid be held stationary, as the capillary tube in Fig. 2(b), the liquid will tend to flow toward the positive electrode.   The reverse effects are also readily pictured, relative motion of the two phases tending to separate the charges, thereby inducing a potential difference.

Electrophoresis.

Consider the movement in an electric field of a charged spherical particle suspended in a stationary liquid.   The frictional drag F_1 on the particle due to its motion through the fluid is given by Stokes' Law,†

$$F_1 = 6\pi\mu r V, \tag{1}$$

where μ is the viscosity of the medium, r the radius, and V the

* A similar diagram can be drawn if the solid be negatively charged.
† Under the conditions considered, the flow of the fluid around the particle is always streamline.

velocity of the particle. The electrical force F_2, acting on a particle carrying a charge C, in an electric field is equal to the product of the charge and the potential gradient dE/dL, *i.e.*,

$$F_2 = C\frac{dE}{dL}. \tag{2}$$

Under steady conditions, which normally develop quickly after imposition of a constant potential, the frictional drag exactly counterbalances the electrical pull, so that equating the two,

$$V = \frac{C\dfrac{dE}{dL}}{6\pi\mu r}. \tag{3}$$

The charged particle under consideration is, in fact, suspended in a dielectric and surrounded at some distance by a cloud of charges (ions) of opposite sign. In effect, the system approximates two charged, concentric, insulated spheres. For such a case the charge is given by the expression,

$$C = \frac{K\zeta rR}{R - r} \tag{4}$$

where K is the dielectric constant of the medium, R the radius of the outer and r of the inner sphere, and ζ the potential difference between them.* This may be written,

$$C = K\zeta r\frac{r + \delta}{\delta} \tag{4a}$$

where δ is the thickness of the dielectric between the surfaces of the spheres. For the special case in which the neutralizing charges in the fluid, external to the particle, are at a distance from the surface of the particle which is large in comparison with its radius (i.e., if δ is large relative to r), this simplifies to

$$C = K\zeta r. \tag{4b}$$

Inserting this last value of the charge in (3), one obtains

$$V = \frac{K\zeta\dfrac{dE}{dL}}{6\pi\mu}, \tag{5}$$

* It should be noted that, whereas E is measured in the direction of motion and is caused by an externally impressed electromotive force, ζ is a self-induced potential difference, measured in a direction at right angles to the particle surface.

an equation frequently quoted and used in interpretation of electrophoretic measurements. Because of the nature of the charges on the particles and around them, it is so highly improbable that δ is large relative to r, that a preferable form of the expression would be

$$V = \frac{K\zeta \frac{dE}{dL}}{6\pi\mu}\left(1 + \frac{r}{\delta}\right) \tag{5a}$$

This indicates that (5) would give values of V which are too low. In fact, the form,

$$V = \frac{K\zeta \cdot \frac{dE}{dL}}{4\pi\mu} \tag{5b}$$

derivable by somewhat similar reasoning, is more generally accepted.

The striking features of (5) and (5b) are that electrophoretic velocity V is independent of particle size and is directly proportional to the applied potential gradient, conclusions which are confirmed by the experimental data.

The fundamental weakness of these formulas lies in the fact that (2) assumes the charged particle isolated, in the sense that the contra-ionic charges in the liquid around it are at a distance too great to influence the magnitude of the force exerted upon it by the external electric field—an assumption which is completely unjustified. Any factor which brings the contra charges closer to the particle surface (see p. 120), *i.e.*, which reduces δ, lessens the effect of the external field upon the particle. This is equivalent to decreasing the magnitude of the charge. Since in the derivation the charge has been assumed proportional to the potential difference ζ between the particle and the liquid around it, any such factor will reduce the computed value of ζ. ζ may be looked upon as a measure of the percentage effectiveness of the particle charge with reference to the external electric field. It is worth emphasizing that the ζ-potential is never subject to direct measurement. It is, however, a most important criterion, even though indirect, of variation in other factors.

The major factor in determining the ζ-potential is the average effective thickness of the Helmholtz double layer, δ. This in turn is determined primarily by the concentration and valence of the contra-ions (see p. 127). That it is indefinite is clear from the diffuse distribution of those ions (see p. 200). The importance of ζ and the justification for its use lie in the fact that, unlike the particle charge, it is substantially independent of particle size, and indeed is found experimentally to be likewise independent of particle shape.*

Electrophoresis can be investigated experimentally by both macroscopic and microscopic means. The former include the moving boundary and transference methods, analogous to the methods of Lodge and Hittorf for the determination of transport numbers and ionic mobilities. In the latter, the rates of travel of single particles under an applied E.M.F. are observed in the microscope.

The rate of movement of the boundary between the sol and the pure liquid is measured conveniently using the apparatus shown in Fig. 3. To secure uniform potential gradient, a solution of the same conductivity as the sol (*e.g.*, potassium chloride) is placed in the U-tube initially and the sol then carefully admitted at the bottom so that a sharp boundary is formed between the liquids in the two limbs.

FIG. 3. Moving Boundary Apparatus.†

Platinum electrodes are inserted and a constant voltage applied. The boundary between the two liquids falls in one limb and

* Efforts have been made to derive the formulas for electrophoresis allowing for the complication indicated. Thus, it has been shown that (5b) should be written:

$$u = \frac{K \zeta \, dE/dL}{4\pi\mu} \cdot f(Mr) \qquad (5c)$$

where M has the dimensions of a reciprocal length and is identified with the reciprocal of the thickness of the ionic atmosphere. It is a function of the ionic strength of the dispersion medium and, to a first approximation, $f(Mr)$ becomes one for any but very dilute solutions and small particles.

† KRUYT, H. R., and VAN DER WILLIGEN, P. C., "Kataphoretischen Messungen," *Koll. Zeit.* 44, 20 (1928). (Theodor Steinkopff.)

rises in the other.   The velocity of rise or fall can be measured readily.   Dividing by the potential gradient dE/dL, gives the electrophoretic velocity or mobility u in centimeters per sec. per volt per centimeter.

The moving boundary method is subject to a number of difficulties and errors.   The sol may be insufficiently colored to allow the boundary to be distinguished.   The overlying liquid frequently exerts an influence, even if the conductivity is the same as that of the sol.   Polarization often occurs at the electrodes, introducing errors in potential gradient.   Travel of the sol boundary downwards is generally greater than that of the moving boundary upwards.   However, many of these difficulties and sources of errors can be surmounted.   Ultraviolet light may be used to render the boundary visible by fluorescence or to follow it by light absorption measurements.   The Tyndall beam may be used for the same purpose.   The influence of the overlying liquid can be largely eliminated by use of the ultrafiltrate from the sol, which has the same ionic constituents and often the same conductivity as the sol itself.   Polarization and back-diffusion of electrolytic by-products can be minimized by the use of reversible electrodes, or by the introduction between the main electrodes of small electrodes upon which is impressed a measured counter E.M.F. just sufficient to prevent current flow through them.   Gravity effects are hard to overcome, but probably tend to counter balance.
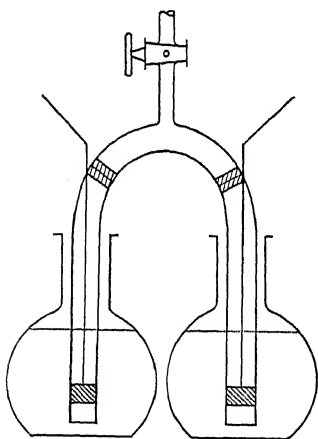
Fig. 4.   Transference Apparatus.*

Transference measurements, although not so simple as the moving boundary method, are capable of giving results of a similar order of accuracy, the limit being governed by the analytical method used.   Two flasks (Fig. 4) are connected by an inverted U-tube, electrodes of platinum foil being wrapped around the lower extremities of the two legs. The sol is placed in each flask, the U-tube filled by suction and a current passed for a known time.   The liquid in the U-tube is then allowed to flow back into the flasks and the change of concentration of the dispersed particles is mea-

* Paine, H. H., *Trans. Faraday Soc.* 24, 412 (1928).

sured.   This may be done by determining some element or group present only in them and not in the dispersion medium.

Consider conditions at any specific cross section of area A in the U-tube where the potential gradient is dE/dL.   Since all particles are moving at constant velocity V, in the absence of convection effects * the concentration M of the suspended particles in mass per unit volume remains unchanged with time. The total mass m of particles transferred in time t is $m = VMAt$. However, the velocity is the particle mobility times the potential gradient, $V = udE/dL$, whence

$$u = \frac{m}{MtA \, \dfrac{dE}{dL}}.$$

Since the current I equals $\kappa A dE/dL$,

$$u = \frac{\kappa m}{MIt}, \tag{6}$$

an equation which has the advantage of eliminating A but requires knowledge of the conductivity κ of the suspension in the U-tube.   When desired, dE/dL can be measured by intermediate electrodes, preferably reversible, carrying negligible current. If the time be short enough so that changes in conductivity in the liquids around the electrodes, induced by concentration changes and electrode reactions, can be neglected and polarization effects ignored, one can take advantage of the fact that the current through the section in question is identical with that through the whole cell.   This, in turn, is equal to the overall voltage times the conductivity of the liquid divided by the effective average length over cross section of the cell.   This ratio, which may be called p, is a function of the shape of the cell, since $I = \kappa A \, dE/dL = \kappa E/p$, where E is the potential difference between the two electrodes.   Substituting in the above equation for I/κ

$$u = \frac{mp}{MEt}, \tag{6a}$$

* Concentrations in the electrode compartments, of course, change with time, but these must not be allowed to penetrate up to the U-tube section in question.

p is determined once for all, preferably using an ordinary electrolyte of known concentration and characteristics.

In the microscopic method the travel of a single particle under the influence of an applied electric field is observed directly, using a small glass cell, which can be mounted on the stage of the microscope. For very small particles the ultramicroscope (p. 133) may be used. When the particles are in Brownian motion, the mean of a number of observations must be made. A difficulty lies in the fact that since the cells used must be small, electroendosmosis occurs along the walls and may even cause the particles to reverse their direction of travel. In a cell so constructed that there is no net motion of the liquid in the direction of the electrodes, granting that under constant potential gradient motion of the particles relative to the liquid is constant, the average velocity of the particles through the cell must equal this absolute velocity relative to the liquid. Hence, survey of the velocity distribution across the cell determines the mobility sought.

Mobilities determined by the different methods agree well. Thus, Table II shows results for four silver hydrosols, using transference and moving boundary methods.

TABLE II. DETERMINATION OF ELECTROPHORETIC MO-
BILITIES BY DIFFERENT METHODS *

Transference, in $cms./sec./volt/cm. \times 10^-$	*Moving Boundary*
37.7	38.5
36.8	38.3
34.7	34.8
35.5	36.6

tially constant. The mobility of suspended particles is of the order of magnitude of the mobilities of the ions. Thus, the mobility of the potassium ion at 21° C. is 64.7×10^{-5}, in comparison with the value of 25.5×10^{-5} for the particles of the silver sol of Table III.

TABLE III. INFLUENCE OF TEMPERATURE UPON THE
ELECTROPHORETIC MOBILITY OF A SILVER SOL *

°C	Mobility	Viscosity, poises	Mobility × Viscosity
3.0	15.1×10^{-5}	16.21×10^{-3}	24.5×10^{-7}
9.9	18.6	13.30	24.7
11.0	19.6	12.82	25.1
21.0	25.5	9.92	25.0
31.0	30.1	7.97	24.0
40.5	37.2	6.58	24.5

Electroendosmosis.

Visualize a capillary tube of radius r, connecting two vessels filled with liquid, an electromotive force being impressed upon two electrodes located in the liquid at the ends of the capillary. A charge will exist on the inner wall of the capillary, counterbalanced by a contra-charge in the liquid itself, located however at an average distance δ from the wall. Because of the electric field through the liquid the charges tend to move. The electric force operating on the charge is equal to the total magnitude of the charge times the potential gradient in the field. Calling C the concentration of charge per unit area of capillary surface and A the total area of the surface, the electrical force equals $CAdE/dL$. Since the force inducing motion operates only in the immediate neighborhood of the surface of the capillary, the liquid will move through the capillary substantially as a rod. In other words, the velocity of motion v of the liquid is constant across the cross section of the capillary, except in the thin layer next the wall between the surface and the contra-charges, *i.e.*, within the Helmholtz double layer itself. Hence, neglecting motion in the thin Helmholtz double layer, the total flow per unit time U is

$$U = \pi r^2 v. \tag{7}$$

* BURTON, E. F., *Physical Properties of Colloidal Solutions*, Ed. 2, Longmans (1921).

This flow is opposed by the friction of slip in the double layer next the wall, which in turn is proportional to the wall area A, and to velocity gradient in the slipping layer v/δ, equalling $\mu A v/\delta$. Equating the opposing forces and eliminating v, one obtains

$$U = \frac{\pi r^2 \delta C}{\mu} \frac{dE}{dL}. \tag{8}$$

Making the Helmholtz assumption that the double layer behaves as a normal plate condenser, so that the potential across it, the ζ-potential, is given by the plate condenser formula, $\zeta = 4\pi\delta C/K$, elimination of δC gives

$$U = \frac{r^2 \zeta K}{4\mu} \frac{dE}{dL}. \tag{9}$$

A similar method may be used to deduce (5b). If one grant that the moving electrophoretic particle is a plate or rod of whatever cross section, provided its major surface is parallel to the direction of motion and end effects are negligible, the fluid frictional resistance to its motion is $\mu A V/\delta$. The electrical force overcoming this friction is $CA\ dE/dL$. Equating the two and making the Helmholtz assumption that the relation between charge on the surface and potential difference between it and the liquid is that of a plate condenser, $\zeta = 4\pi\delta C/K$, one obtains

$$V = \frac{K\zeta}{4\mu\pi} \frac{dE}{dL}. \tag{5b}$$

Because it is desirable to study electroendosmotic behavior of many materials which it is difficult to secure in the form of capillaries, it is important to know the characteristics of such flow through the irregular pores of solid masses of particles, such as cellulose, sintered glass, etc. Assuming that the voids of such a mass can be considered as equivalent to a number of capillaries of effective radius r and length L, the electrical resistance through one of these capillaries is $R = L/\kappa A = L/\kappa\pi r^2$. Since the voltage drop is equal to the product of current times resistance, it follows that $E/L = I/\kappa\pi r^2$, whence

$$U = \frac{\zeta I K}{4\pi\mu\kappa} \tag{9a}$$

In this equation all quantities except the ζ-potential are either subject to direct measurement or else are properties of the liquid

in the capillaries. If one is willing to grant that the capillaries have no effect on the properties of the liquid, the equation makes it possible to determine the ζ-potential directly. The permissibility of this assumption is certainly open to question. The conductivity can be measured directly within the capillaries themselves by the use of an alternating current which would presumably eliminate the effects of flow. Furthermore, as will be seen from the derivation, the conductivity in question is the average of all the liquid in the pores. This is not true of the viscosity and dielectric constant, both of which refer to the properties of the Helmholtz double layer alone, *i.e.*, of the liquid in the immediate neighborhood of the solid surface past which it is flowing. No satisfactory method for determining the effects of the capillary on either of these has been developed.*

Figure 5 shows a typical apparatus for the study of electroendosmosis. The porous solid to be studied constitutes the
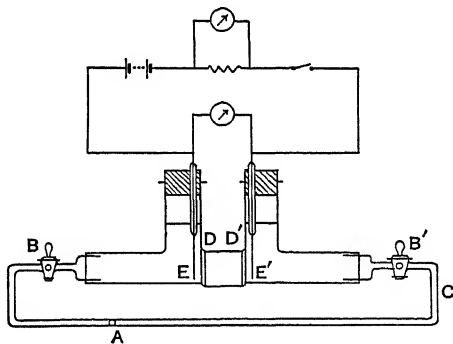
FIG. 5.   Electroendosmosis Apparatus.†

diaphragm *DD'* between the electrodes *EE'*. By suitable manipulation of the three-way stopcocks *BB'* an air bubble *A* can be introduced into the capillary tube *C* to indicate the flow of the liquid. The results show that the volume endosmosed in a given time is proportional to the applied E.M.F. and varies with the viscosity (temperature), as indicated by (9). Table

---

* Thus, while the work of Bowden indicates that the effect of capillaries on viscosity is negligible down to extremely small dimensions, his method gives an average value of the viscosity for the capillary as a whole and not for the double layer alone.

† HEPBURN, H. C., *Proc. Phys. Soc.* **38**, 363 (1926). Based on design by T. R. Briggs, *J. Phys. Chem.* **22**, 256 (1918).

IV gives the time T required for a bubble to move over 5 cms. of the capillary tube C for different applied voltages E. The reciprocal of their product should, of course, be constant.

TABLE IV.   EFFECT OF APPLIED POTENTIAL DIFFERENCE ON ELECTROENDOSMOSIS *

E, volts	T, sec.	$1/TE \times 10^4$
8.24	299.0	4.06
12.34	200.8	4.04
16.48	150.4	4.03
20.62	119.7	4.05

Electroendosmosis is profoundly influenced by the polarity of the liquid. The results in Table V, obtained by Fairbrother and Balkin working with very carefully purified liquids and a sintered-glass diaphragm, show marked parallelism between electroösmotic flow and polarity as indicated by dielectric constant and dipole moment (pp. 15–16).

TABLE V.   ELECTROÖSMOSIS OF ORGANIC LIQUIDS †

Liquid	$V \times 10^5$ cc./sec./volt ^{a}	Dielectric Constant	Dipole Moment $\times 10^{18}$
Formamide	5.68	84	?
Water	16.16	81	1.7 c.s.u.
Nitrobenzene	5.53	36	3.9
n-Propyl alcohol	1.95	22	1.75
Aniline	0.15	7.2	1.51
Chloroform	0.63	5.2	——
Ether	−0.96	4.36	1.22
Benzene	0	2.26	0
Carbon tetrachloride	0	2.23	0

^{a} Minus sign indicates flow toward the cathode. The values of velocity are relative only.

Streaming Potential.

In view of the fact that a potential gradient through a capillary tube or structure can induce flow of liquid through it, one would expect that if liquid flow through the capillary be caused by an applied mechanical pressure, a corresponding electromotive force would be generated. This is actually the case, the magnitude of the effect paralleling the polarity of the

* HEPBURN, H. C., *Proc. Phys. Soc.* **38**, 363 (1926).
† FAIRBROTHER, F., and BALKIN, M., *J. Chem. Soc.* 389 (1931).

liquid, as in electroendosmosis. Thus, the normal aliphatic alcohols flowing through a porous cellulose diaphragm, under fixed conditions of diaphragm structure and pressure gradient, give streaming potentials which decrease by 36 millivolts for each additional methylene group introduced into the hydrocarbon chain.* Similarly, while benzene flowing through the same diaphragm shows no potential, the introduction into its molecule of substituent radicals develops a potential increasing progressively in the order $CH_3 < Cl < Br < NH_2 < NO_2$.

Efforts have been made to compute ζ-potentials from streaming-potential measurements and to compare them directly with electroendosmotic effects in the same mechanical set-up. Unfortunately, however, the computation methods employed are subject to serious question on theoretical grounds and only in the case of specific solutions have apparent checks been obtained.

Migration Potential.

If one can force suspended particles to move through a liquid, a phenomenon which is the reverse of electrophoresis should develop, *i.e.*, an E.M.F. should be generated. Gravity and centrifugal force are available for movement of the particles and qualitative measurements have demonstrated the development of a potential due to this factor.

The Double Layer.

Colloidal particles may become charged either by adsorption of ions from the dispersion medium or by their own ionization. The former mechanism characterizes many suspensoids; *e.g.*, the particles in a gold sol formed by the Bredig method using dilute hydrochloric acid become negatively charged with chloride ions adsorbed from the solution. Acid, basic, or amphoteric materials, on the other hand, when in colloidal solution, usually charge themselves by their own ionization. Thus the proteins (see p. 161) are amphoteric and in acid solution give positively charged particles, while in alkaline solution the particles carry a negative charge. Such materials, are, in fact, colloidal electro-

* MARTIN, W. M., and GORTNER, R. A., *J. Phys. Chem.* 34, 1509 (1930).

lytes, one of the ions having colloidal dimensions. Whatever the source of the charge, the colloidal particle will be surrounded in the solution by a layer of ions of opposite sign, thus giving, with its own ions, the so-called double layer.

Helmholtz conceived the double layer as made up of oppositely charged layers at a definite distance, of the order of molecular dimensions, apart. Thus, a very sharp break in the electric potential would occur across the double layer, this potential difference being the ζ-potential. Thermal agitation makes such a rigid double layer unlikely. The outer ions must form a diffuse layer, their concentration gradually shading off until it is equal to their concentration in the main bulk of the solution. The arrangement of ions will be governed by a distribution law similar to that of Fig. 1 (p. 1) so that the double layer will lack the definiteness of the Helmholtz concept, resembling Fig. 5 (p. 119) rather than Fig. 2(a) of this chapter.

Electrolytes and Electrokinetic Phenomena.

When an electrolyte is added in sufficient quantity to a suspensoid, the mobility of the particles in an electric field decreases, and at a certain point precipitation occurs. The ζ-potential calculated from the mobility at this point is called the "critical potential." The value of the critical potential is, in most cases, not zero, but differs considerably for different substances.* For a given suspensoid it is characteristic of the point at which instability develops. Table VI shows results obtained by Loeb for an aqueous collodion dispersion prepared according to the method described on page 178. The ζ-potentials were determined by electrophoretic measurements and are reproducible to ± 2 millivolts. It is clear that there is a critical potential approximating 13 millivolts, independent of the precipitating electrolyte. Similarly, for gold sols the critical potential is found to be nearly zero and for oil drops suspended in water it is about 40 millivolts. Thus, having once determined the critical potential of a given suspensoid, its stability can be forecast from its electrokinetic potential.

* When the particles of a sol are strongly solvated, as in the proteins and some clays, the sol may still be stable when it shows no motion in an electric field, *i.e.*, its electrokinetic potential is zero.

Electrolytes also influence electroendosmosis.   In pure water the charge on a solid capillary tube or diaphragm depends largely upon its character.   Basic materials, such as zinc oxide, tend

TABLE VI. ζ-POTENTIAL AND STABILITY OF COLLODION DISPERSIONS AT pH 5.8 *

Electrolyte	Minimum Concentration Required for Precipitation	ζ, in Millivolts	Maximum Concentration at which Suspension is Stable	ζ, in Millivolts
LiCl	M/2	10	M/4	17
NaCl	M/2	10	M/4	14
KCl	M/4	14	M/8	21
Na$_2$SO$_4$	M/4	13	M/8	19
Na$_4$Fe(CN)$_6$	M/16	13	M/32	21
MgCl$_2$	M/16	11	M/32	15
MgSO$_4$	M/16	15	M/32	19
CaCl$_2$	M/32	14	M/64	17
LaCl$_3$	M/2048	14	M/4096	21

to form their own electric double layer, presumably by the loss of hydroxyl groups into the solution, leaving the diaphragm positively charged.

$$ZnO \rightarrow Zn(OH)_2 \rightarrow Zn(OH)^+ + (OH)^-$$
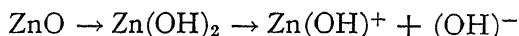
Diaphragms of acidic materials, on the other hand, become negatively charged.   Most neutral substances, possibly because of the more ready adsorbability of the hydroxyl ion, tend to possess a negative charge in water.   Thus, sulfur, glass, naphthalene, carborundum, cotton, silica, and tungstic acid all bear negative charges in water, while alumina, barium carbonate, iron oxide, zinc oxide, and magnesia are charged positively.   As in the case of electrophoresis, the ζ-potential may be depressed by addition of electrolytes; the velocity of electroösmosis is at first reduced and then reversed.

The valence effect (p. 125) is also marked, as shown in Table VII for endosmosis through a glass capillary.   Far smaller concentrations of di- and trivalent than of monovalent ions are required to reduce the endosmotic flow by a given amount. Organic ions also have a powerful effect and a peculiar phenom-

* LOEB, J., *J. Gen. Physiol.* 5, 109 (1922).

enon is noticed with caustic soda, the velocity of flow being actually increased.   Probably the hydroxyl ion is powerfully adsorbed and overcomes the discharging effect of the sodium

TABLE VII.   THE EFFECT OF ELECTROLYTES UPON ELEC-
TROÖSMOSIS THROUGH A GLASS CAPILLARY *

Conc. in millimols $\times$ 10^{-3}/liter	Relative Velocity of Endosmosis toward Cathode						
	0	5	10	20	50	225	4000
NaCl	50	48	47	43	38	26	5
1/2 K₂SO₄	50	47	44	41	38	26	8
Crystal violet	50	50	42	29	4	+	
BaCl₂	50	44	37	29	24	——	——
Ce(NO₃)₃	50	33	27	14	9	3	——
Th(NO₃)₄	50	4	+	——	——	——	——
NaOH [a]	50	——	99	99	——	——	——

[a] Quartz capillary.
+ Indicates flow towards anode.

ion.   The resemblance to the coagulation of suspensoids is also seen in Table VIII, where the liminal concentrations γ required to precipitate a gold sol are compared with the concentrations

TABLE VIII.   THE ACTION OF ELECTROLYTES UPON
ELECTROÖSMOSIS AND SUSPENSOIDS *

Electrolyte	c, in millimols $\times$ 10^{-3}/liter	γ, in millimols/liter
NaCl	260.0	24
K₂SO₄/2	270.0	23
HNO₃	39.0	10
AgNO₃	47.0	——
Morphine chloride	51.0	0.54
Neufuchsin	41.0	0.002
K-Benzoate	29.0	17
BaCl₂	37.0	0.35
HgCl₂	3.8	0.014
Al₂(SO₄)₃/2	4.5	0.009
Ce(NO₃)₃	8.6	0.003
Th(NO₃)₄	2.7	

c necessary to reduce the volume endosmosed in a given time by one half.   Salts of the univalent alkaline metals are least effective in both cases, salts of the di- and trivalent metals being

* ELISSAFOFF, G. V., Z. physik. Chem. 79, 385 (1912).

considerably more so.   The organic and heavy metal ions occupy an anomalous position.

Streaming potential is also markedly affected by electrolytes, the ion of sign opposite to that attached to the solid wall again being the active agent.   Figure 6 shows the streaming
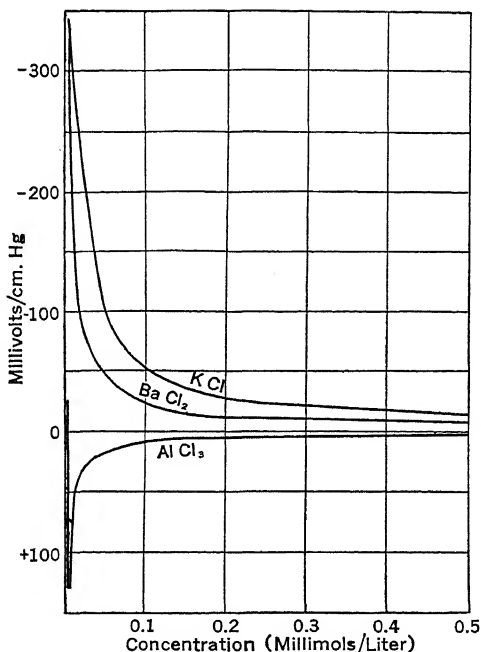
FIG. 6.   Streaming Potentials through a Glass Capillary.*

potentials for a number of electrolytes with a glass capillary, determined by Kruyt, the potential expressed as millivolts per cm. of mercury pressure applied to the liquid.   Aluminum ions reverse the sign of the potential at surprisingly small concentrations.   At higher concentrations the chlorine ion begins to exert its discharging action on the double layer, giving an "irregular series" for streaming potential, as in the coagulation of suspensoids (p. 131).   In other words, the behavior of the double layer is independent of the relative motion of the solid and liquid.

* KRUYT, H. R., "Strömungspotentiale Kolloidstabilitate," *Kolloid Z.* **22**, 81 (1918). (Theodor Steinkopff.)

Monaghan, White, and Urban have determined the concentrations of several salts at which the electroendosmosis velocity, the streaming potential, and electrophoretic velocity of Pyrex glass are all zero, *i.e.*, the iso-
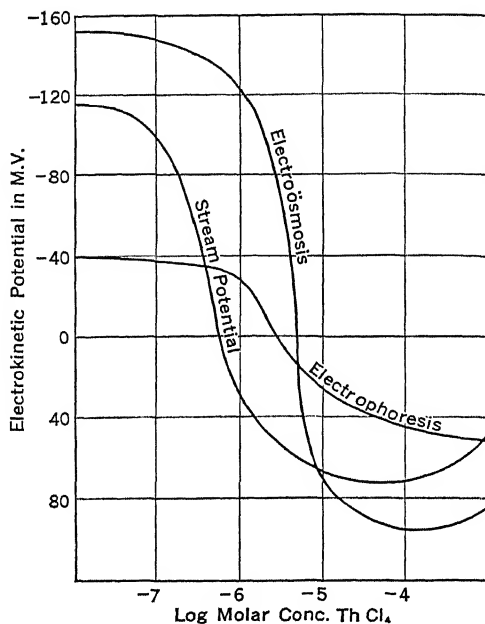
FIG. 7.   Effect of $ThCl_4$ on ζ-Potentials as Determined by
Various Methods of Measurement.*

electric concentrations.   Their results are shown in Table IX, while Fig. 7 illustrates the varying values of the ζ-potential obtained over a range of

TABLE IX.   ISOELECTRIC CONCENTRATIONS FOR GLASS AND
PROTEIN COATED GLASS *

	Isoelectric pH for Gelatin Coated Glass	*Isoelectric Concentration for $ThCl_4$ with Pyrex*	*Isoelectric Concentration for $AlCl_3$ with Pyrex*	*Isoelectric Concentration for $FeCl_3$ with Pyrex*
Stream potential	4.75	4×10^{-7} M	1×10^{-6} M	1×10^{-6} M
Electroösmosis	4.75	3×10^{-6}	3×10^{-6}	3×10^{-6}
Electrophoresis	4.75	3×10^{-6}	3×10^{-6}	3×10^{-6}

concentrations for one salt, thorium tetrachloride.   For protein-coated glass (a film of gelatin being readily formed by adsorption), the hydrogen-ion

* MONAGHAN, B., WHITE, H. L., and URBAN, F., *J. Phys. Chem.* 39, 585 (1935).   (The Williams and Wilkins Co.)

concentrations at which the three phenomena were zero, were always the same.

Colloidal Electrolytes.

Reference has already been made to the fact that many colloids obtain their stabilizing charge by self-ionization and may be called colloidal electrolytes. Soaps, clays, gum agar, many dyestuffs (*e.g.*, Congo Red) may be instanced as belonging to this type. Clay, for example, may be regarded as a mixed alkali-alkaline earth metal salt of an alumina-silica complex of high molecular weight. By treatment with acid and subsequent washing or by electrodialysis the alkali and alkaline earth metals
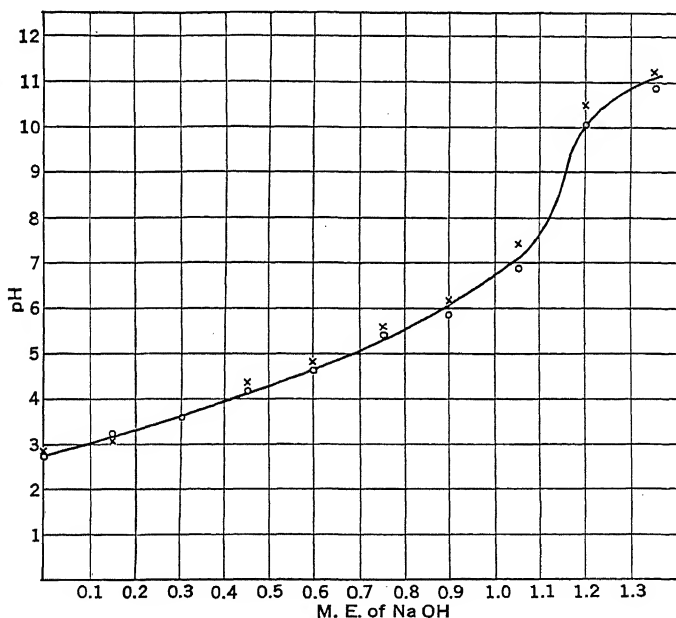
Fɪɢ. 8.   Titration of Freshly Electrodialyzed Ultrabentonite.*
Basis: 1 gram oven dried clay.

can be removed and a hydrogen clay obtained. This behaves as an acid,

$$HB + NaOH \rightarrow Na^+ + B^- + H_2O,$$

Fig. 8 showing the electrometric titration curve for hydrogen

* Tʜᴏᴍᴘsᴏɴ, W. I., *Mass. Inst. Tech. Thesis* (1937).

bentonite (p. 240) determined with the aid of the glass electrode. From the break in the curve at pH = 9 it is possible to calculate the equivalent weight of the clay as approximately 870. Nevertheless, the clay particles are colloidal in size, are visible under the ultramicroscope, and under suitable conditions can be made to gel (p. 222). In an electric field they also travel to the anode, indicating their negative charge. In other words, the clay behaves as a typical colloid.

Light is thrown on the behavior of colloidal electrolytes by study of their electrical conductivity. For example, if a known quantity of hydrochloric acid be added to a dilute solution of pure gelatin, the hydrogen-ion concentration in the solution can be measured with the hydrogen electrode. The acid combined with the gelatin reacting as a base equals the difference between the hydrogen ion experimentally determined in the solution and the known quantity of acid added to it. If now the conductivity of such a solution be measured experimentally, that portion of the conductivity which is due to the mobility of the hydrogen and chlorine ions can be calculated from their known mobility. The excess conductivity must be that of the gelatin constituent itself. Data on such a case are presented in Table X, all measurements being made in solutions which contained the same total quantity of pure gelatin per unit volume. The concentration of gelatin chloride reported in the table is the equivalent of the chlorine combined with gelatin, obtained, as indicated above, by difference between the total chlorine added as HCl and the experimentally measured hydrogen-ion concentration. Inspection of the table shows a well-defined maximum in the equivalent conductivity of gelatin chloride thus computed. At low acid concentrations the gelatin particle carries little charge, but, because of its size, exhibits a large frictional (Stokes' Law) resistance to movement through the solution. It has, therefore, low mobility. As acidity increases, the number of charges on the individual gelatin ion increases rapidly. This induces a marked additional electrical pull with presumably relatively little effect on the friction. When, however, acid concentration gets high, the table shows that the mobility of the gelatin particle decreases. This may be due to swelling of the particle in

strong acid solution with resulting increase in friction, or to incomplete ionization in the relatively strong solutions in question. Somewhat similar behavior has been noted with soap

TABLE X. EQUIVALENT CONDUCTIVITIES AT 25° C. OF
GELATIN CHLORIDE

Concentration equivalents/liter × 10⁴	Conductivity × 10⁻⁴	Equivalent Conductivity
6.75	0.484	71.6
14.1	1.051	74.5
23.9	1.75	73.2
36.5	2.65	72.7
47.9	3.57	74.7
55.8	4.585	82.1
63.0	5.405	86.0
69.7	5.99	86.0
74.0	6.07	82.1
79.2	6.21	78.5
82.9	6.25	75.4
85.2	5.37	63.2

solutions, such as potassium palmitate, over a portion of the conductivity concentration curve, conductivity increasing with rising concentration. In this case it appears probable that aggregates of neutral soap molecules and fatty acid ions are formed in greater numbers as the concentration rises, but the increased frictional resistance is more than offset by the increased charge on the aggregates. With gelatin, particle size presumably remains constant but more and more amino groups on the particle are ionized as the acid concentration is increased.

With amphoteric colloidal electrolytes, e.g., the proteins (p. 159) in the neighborhood of neutrality, phenomena are still more complicated since ionization either as acid or base can occur.

Isoelectric Point.

The behavior of amphoteric colloids towards acids and bases is approached through analogy to the behavior of simple univalent ampholytes. Consider, for example, a material such as glycine $CH_2.NH_2.COOH$. Its dissociation as an acid is governed by the mass action equation:

$$k_a = \frac{[H^+][CH_2.NH_2.COO^-]}{[CH_2.NH_2.COOH]}. \tag{10}$$

Its corresponding dissociation as a base is given by

$$k_b = \frac{[CH_2.COOH.NH_3^+][OH^-]}{[CH_2.NH_2.COOH]}. \tag{11}$$

In a specific solution of glycine the material will dissociate in both these ways, but inasmuch as the concentration of the completely undissociated glycine appears in both equations it can be eliminated from them, giving

$$\frac{[H^+][CH_2.NH_2.COO^-]}{k_a} = \frac{[CH_2.COOH.NH_3^+][OH^-]}{k_b}. \tag{12}$$

Defining the isoelectric point as that at which the number of positive charges on the glycine is equal to the number of negative ones, at this point the concentration of negative glycine ions must equal that of positively charged ones. Thus, remembering that $k_w = [H^+][OH^-]$, (12) reduces to

$$[H_i^+] = \sqrt{\frac{k_a k_w}{k_b}} \tag{13}$$

where $[H_i^+]$ is the hydrogen-ion concentration at the isoelectric point. This equation assumes a uni-univalent ampholyte, the dissociation of which obeys the Mass Action Law.*

When one is dealing with a polyvalent ampholyte, which can dissociate both as an acid and as a base in a number of ways, the corresponding acid and basic constants being k_{a1}, k_{a2}, $\cdots$ k_{an} and k_{b1}, k_{b2}, $\cdots$ k_{bm}, respectively, again applying the Mass Action Law for all stages of dissociation on the assumption that ionization of any specific group is completely unaffected by the simultaneous ionization of the other groups, the hydrogen-ion concen-

---

* There is much evidence to indicate that in pure solution amphoteric electrolytes of this type may exist as a kind of internally neutralized ion, the structure

$$R \overset{\displaystyle COO^-}{\underset{\displaystyle NH_3^+}{<}}$$
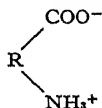
being called a zwitterion. Thus, at the isoelectric point, the particle carries no net electric charge but, in contrast to the older ideas given above, may be ionized in high degree. The above derivation ignores the possible development of zwitterions, but it is readily shown that the equation is still valid provided all stages of ionization follow the Mass Action Law.

tration at the isoelectric point is given by the equation

$$[H_i^+] = \sqrt{\frac{k_{a1} + k_{a2} + \quad k_{an}}{k_{b1} + k_{b2} + \quad \mathit{v}_{bm}}} \tag{13a}$$

The character of the expressions for the hydrogen-ion concentration at the isoelectric point makes it clear that its value is independent of the concentration of ampholyte and added electrolyte as well. On the other hand, the hydrogen-ion concentration of a solution of the pure ampholyte uncontaminated with other electrolytes and uncombined with either acid or base is, in general, not identical with that at the isoelectric point. Because the sum of the charges on anions and cations in a given solution must be equal, one may write

$$[H^+] + \frac{k_b[CH_2 . NH_2 . COOH]}{[OH^-]}$$
$$= [OH^-] + \frac{k_a[CH_2 . NH_2 . COOH]}{[H^+]}$$

whence

$$[H^+]^2 = \frac{k_w + k_a[CH_2 . NH_2 . COOH]}{1 + \dfrac{k_b}{k_w} [CH_2 . NH_2 . COOH]}. \tag{14}$$

It follows from this last expression that the hydrogen-ion concentration in a solution of a pure ampholyte is a function of its concentration, in contradistinction to its isoelectric point.

A characteristic of the isoelectric point is that the total ionization of the ampholyte is a minimum. Again taking glycine as example, if x be the fraction of dissociated electrolyte,

$$x = \frac{[CH_2 . NH_2 . COO^-] + [CH_2 . COOH . NH_3^+]}{[CH_2 . NH_2 . COOH]} \tag{15}$$
$$= \frac{\dfrac{k_a}{[H^+]} + \dfrac{k_b}{[OH^-]}}{\dfrac{k_a}{[H^+]} + \dfrac{k_b[H^+]}{k_w}}$$

Differentiating with respect to the hydrogen-ion concentration

$$\frac{dx}{d[H^+]} = \quad \frac{k_a}{[H^+]^2} + \frac{k_b}{k_w} \tag{16}$$

the condition for minimum ionization can be found by equating $dx/d[\text{H}^+]$ to zero, whence:

$$[\text{H}^+]_{\text{min.}} = \sqrt{\frac{k_a k_w}{k_b}} = [\text{H}^+]_i. \qquad (17)$$

The trouble with these equations is that there is much evidence to indicate that the Mass Action Law does not apply to the dissociation of ionizable materials of colloidal size. Nevertheless, while these equations are not quantitatively applicable to colloids, e.g., to protein sols, the qualitative conclusions derivable from them are apparently valid for all colloidal ampholytes. Thus, all such materials possess an isoelectric point, at which the substance shows no movement in an electric field. Furthermore, this point is little, if at all, affected by the concentration of the colloid.

The most dependable method for the determination of the isoelectric point of an amphoteric colloid is the determination of its electrophoretic movement in solutions of varying pH, for which the apparatus of Fig. 3 may be used. The isoelectric hydrogen-ion concentration is that at which motion reverses. In this way, the isoelectric points of many amphoteric substances have been determined, a selection being shown in Table XI.

TABLE XI.   ISOELECTRIC POINTS OF AMPHOTERIC
SUBSTANCES *

Egg albumin	4.55	Pepsin	2.5–3.3
Casein (cow's milk)	4.7	Sericin	4.2
Gelatin	4.7	Thyroglobulin	4.6
Collagen (hide powder)	5.0	CO-Hemoglobin	7.1
Wool	4.9	Fibroin	5.0

As might be predicted from the foregoing, hydrogen-ion concentration has an important influence on the physical properties of amphoteric colloids. Because the charges on the colloid are a minimum at the isoelectric point, its tendency to go into solution and its stability when dissolved are likewise a minimum. This fact exhibits itself in the swelling of such materials when treated with a solvent under conditions (e.g., of

* In many cases, the precise values of the isoelectric points are in dispute or cover a range; those listed here should be regarded only as approximate.

temperature) such that complete solution does not occur. The character of the results is shown in Fig. 9 for gelatin in monobasic acids and alkalies. There is a minimum in the swelling curve at the isoelectric point, swelling increasing rapidly as one diverges from it in either direction by addition of acid or alkali. However, each branch of the curve goes through a maximum, perhaps due to a desolvating effect of the added electrolyte as its concentration becomes high (see also discussion on p. 216).

The affinity of the colloid for the solvent as influenced by the hydrogen-ion concentration of the latter can be approached from another angle. Thus, if to the solution of the colloid one add a neutral liquid, which is itself not a solvent for the colloid but is miscible in all proportions with the original solvent, the amount of this added material necessary to initiate precipitation is a measure of the stability of the solution. This amount is a minimum at the isoelectric point, and, if the quantity be plotted against pH, it gives a curve paralleling that of Fig. 9.
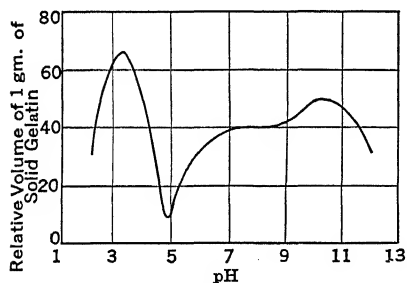
FIG. 9.   The Swelling of Gelatin.*   FIG. 10.   The Viscosity of
                                         Gelatin Solutions.*
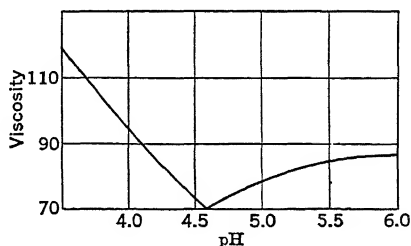
The physical properties of the solution show similar influences of pH. Figure 10, due to Loeb, shows the viscosity of a solution of constant gelatin concentration per unit volume adjusted to various values of pH. The flow interference of the gelatin particles is clearly a minimum at the isoelectric point and increases as the colloid combines with either acid or base. Whether the cause of the decreased flow is merely increased solvation or is also influenced by the dragging effect of contra-ions external to the gelatin particles themselves is not entirely clear. Figures

* LOEB, J., *Proteins and the Theory of Colloidal Behavior*, McGraw-Hill (1922).

11 and 12, also due to Loeb, show the corresponding effects of pH on osmotic pressure (see p. 216) and on the electric conductivity. The latter is obviously confused by the mobility of the ions of the inorganic electrolyte, the presence of which is essential to modify the pH.
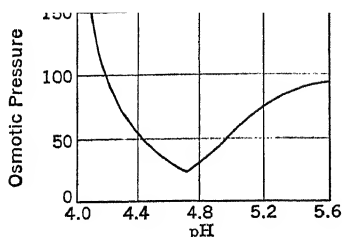
FIG. 11.   The Osmotic Pressure
of Gelatin Solutions.*
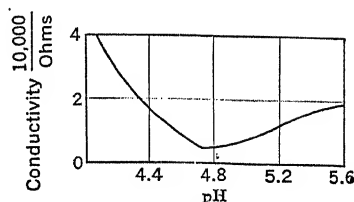
FIG. 12.   The Electrical
Conductivity of Gelatin Solutions.*

The stronger the acidic and basic groups of the ampholyte, the more sharply defined are the changes in properties of the solution in the neighborhood of the isoelectric point. If both groups are very weak, the material ceases to act as either acid or base unless in the presence of high concentrations of strong acids. In other words, the material becomes in effect neutral and the changes in properties of its solution in the neighborhood of its isoelectric point are much less marked.

Donnan Equilibrium.

An important consequence of the colloidal size of one of the ions of a colloidal electrolyte is the behavior of such electrolytes when confined by a membrane, permeable to electrolytes but non-permeable to colloids. Visualize, for example, the cell illustrated in Fig. 13, containing a colloidal electrolyte, such as gelatin. The cell is separated into two portions A and B by a membrane permeable to normal electrolytes but impermeable to gelatin or its ions. The space A initially contains a solution of pure gelatin in water, and to B is added a dilute, aqueous solution of hydrochloric acid. The acid will diffuse through the membrane and combine to a greater or less degree with the gelatin. This diffusion is allowed to proceed until equilibrium is attained.

*LOEB, J., *Proteins and the Theory of Colloidal Behavior*, McGraw-Hill (1922).

The hydrochloric acid initially diffusing into the gelatin compartment will be almost completely neutralized and converted into gelatin chloride. Because of this chemical reaction, the acid concentration in A will be kept low, and diffusion of hydrochloric acid from B into A will continue. It can readily be demonstrated in the laboratory that diffusion continues until the total chloride ion concentration in the gelatin compartment far exceeds that on the other side of the membrane.

Assume for the moment that the ionization of the gelatin chloride can be considered substantially complete, *i.e.*, that the activity of the chloride ion in the gelatin compartment is equal to its total concentration. Because, when equilibrium is attained, the chloride-ion concentration in compartment A is so high, clearly its tendency to diffuse back into B must be correspondingly great. Since this diffusion does not occur in practice, there must be some counterbalancing effect. This is the diffusing tendency of the hydrogen ion. Experimental measurement of the hydrogen-ion concentration on the two sides of the membrane shows it to be far higher on the acid side than in the gelatin compartment. Neither chloride nor hydrogen ion can diffuse without dragging the other ion with it, on account of their mutual electrostatic attraction. Movement of the two ions through the membrane can only occur when the two strike it simultaneously. The frequency of such collisions will be proportional to the product of the concentrations of the two ions on the same side of the membrane. At equilibrium the same number of hydrogen and chloride ions pass from A to B as from B to A; hence the product of their concentrations on the two sides of the membrane must be equal:

$$[H^+]_A[Cl^-]_A = [H^+]_B[Cl^-]_B. \tag{18}$$

This relation, the so-called Donnan membrane equilibrium, has repeatedly been confirmed experimentally. Table XII gives results of a series of experiments on a 1% solution of gelatin with

Semipermeable
Membrane

Gelatin in | Dilute HCl
Water

(A) | (B)

FIG. 13. Donnan Equilibrium Cell.

varying amounts of hydrochloric acid, inner and outer solutions being separated by a collodion membrane. The chloride-ion concentration was determined by titrating the two solutions with sodium hydroxide. The hydrogen ion was measured po-

TABLE XII.   MEMBRANE EQUILIBRIUM WITH GELATIN AND HYDROCHLORIC ACID *

pH of gelatin solution at equilibrium	$\log \dfrac{[H^+]_B}{[H^+]_A}$	$\log \dfrac{[Cl^-]_A}{[Cl^-]_B}$
4.13	0.56	0.48
3.69	0.58	0.51
3.30	0.50	0.59
3.10	0.49	0.44
2.92	0.44	0.44
2.78	0.44	0.38
2.46	0.33	0.35
2.26	0.23	0.22
2.01	0.15	0.15
1.76	0.10	0.11

tentiometrically with a hydrogen electrode. The table shows that, within the experimental accuracy, equation (18) holds good. It will be noted that the maximum deviations correspond to concentration differences of about 20%. Inasmuch as the two solutions on opposite sides of the semipermeable membrane have widely different concentrations of the active diffusible ions, if one will set up an electrolytic cell with identical suitable electrodes immersed in the two solutions, there will be developed the usual concentration cell potential, the value of which is given by the conventional equation:

$$\Delta E = \frac{RT}{F} \log_e \frac{[H^+]_B}{[H^+]_A} = \frac{RT}{F} \log_e \frac{[Cl^-]_A}{[Cl^-]_B}. \qquad (19)$$

Loeb has measured the membrane potential, ΔE, for the gelatin-hydrochloric acid system, using the cell

| calomel electrode | sat'd KCl | external solution | internal solution | sat'd KCl | calomel electrode |
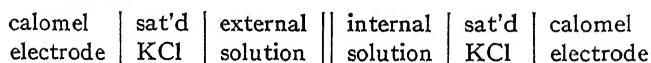
The measured potentials were found to be in close agreement with those calculated from the hydrogen-ion concentrations determined inside and outside the membrane with the aid of the hydrogen electrode.

* LOEB, J., *J. Gen. Physiol.* **3**, 667 (1921).

Osmotic Pressure of Colloidal Electrolytes.

The system just discussed, consisting of the salt of a colloidal base separated by a membrane impermeable to the base from a solution of the corresponding acid, is not at equilibrium. Even granting that the molecular weight of the colloid is such that its osmotic pressure is negligible, the anions combined with the colloid are not free to move through the membrane, despite the fact that it is permeable to them. This is because of the electrostatic attraction of those anions for the positive charges on the colloidal base. The diaphragm is, therefore, in practical effect impermeable to the anions. In consequence, the latter exert an osmotic pressure at the face of the diaphragm, tending to draw solvent through it and dilute the colloidal solution within. It is obvious that the osmotic pressure within the membrane is partially offset by that outside, and the net, experimentally observable pressure is the difference between the two. Granting that the solutions are sufficiently dilute so that the osmotic pressures follow the gas laws, one can compute the osmotic effect, using the total concentrations of solute species, whether diffusible or not.

Again considering a gelatin-hydrochloric acid system, let x be the concentration of hydrogen and chlorine ions in the solution outside the membrane, y and $(y + z)$ their concentrations, respectively, inside the membrane, and t the total concentration of gelatin within the membrane. Then, if e is the excess of diffusible ions within the membrane,

$$e = 2y + z - 2x,$$

and the osmotic pressure P is given by

$$P = RT(t + 2y + z - 2x). \tag{20}$$

In dilute solutions, however, t is small and to a first approximation

$$P = RT(2y + z - 2x). \tag{21}$$

However, from the Donnan equilibrium,

$$x^2 = y(y + z). \tag{22}$$

Hence, substituting for z,

$$P = RT \frac{(x - y)^2}{} \tag{23}$$

P, x, and y are all susceptible to direct measurement, hydrogen-ion concentrations being readily determined with the hydrogen electrode.   In Fig. 14 the osmotic pressure experimentally ob-
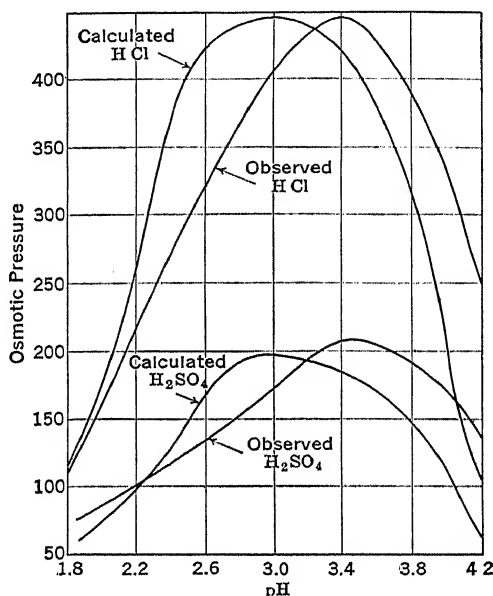
FIG. 14.   Calculated and Observed Osmotic Pressures
of Gelatin-Acid Solutions.*

served by Loeb is plotted along with the osmotic pressure calculated according to (20).   The shapes of the curves are very similar, the slightly higher maximum of the observed osmotic pressure probably indicating that the concentration of the protein itself, t, cannot be entirely neglected.   Furthermore, all monobasic acids, independent of their structure, are found to give the same osmotic pressure curve with varying hydrogen-ion concentration.

* LOEB, J., *Proteins and the Theory of Colloidal Behavior*, McGraw-Hill (1922).

Dibasic acids give a lower maximum osmotic pressure, equations corresponding to (21) and (22) being easily developed. Figure 14 shows the agreement obtained between calculated and observed values.

The shape of the osmotic pressure curve can be deduced by rewriting (22) in a somewhat different form. From the membrane equilibrium,

Substituting this in (20),

$$P = (\sqrt{4y^2 + 4yz + z^2} - \sqrt{4y^2 + 4yz})RT. \tag{24}$$

At the isoelectric point where no acid has been added $z = 0$ and $P = 0$. As acid is added, z at first increases more rapidly than y until the acid combining power of the protein is satisfied. Thereafter, z rapidly becomes negligible in comparison to y so that P once more tends to zero, or the value of P passes through a maximum, tending to zero either when no acid or a large excess of acid is present.

The addition of neutral salts to protein solutions is also often of importance. If sodium chloride be added to the gelatin-hydrochloric acid system, clearly z will be unchanged but y must increase. From (24) it follows that the osmotic pressure diminishes, a result confirmed experimentally. Addition of salts with no common ion, $e.g.$, sodium nitrate, also has the same effect. It may be assumed that the gelatin chloride is partially converted to gelatin nitrate, but z still has the same value and the total concentration of diffusible ions is increased.

Gels require no external membrane, the surface of the gel itself in contact with dilute acid serving in this capacity. Proctor and Wilson * showed not only that the chloride-ion concentration in the gelatin jelly in excess dilute HCl was considerably greater than that in the external solution, but that the total excess of freely diffusible ions e is given quantitatively by

$$e = 2y + z - 2x.$$

In order to equalize this excess, water diffuses into the pores and open spaces of the gel, causing it to swell until the diffusing force is just balanced by the cohesive forces of the jelly structure. The pull into solution and hence the swelling will be proportional to e. The maximum swelling is to be expected, therefore, at the same hydrogen-ion concentration as the maximum osmotic pressure. Inspection of Figs. 9 and 14 shows this to be the case, the shape of the curves being similar, with both maxima occurring at approximately the same pH, namely, 3.3. Efforts have been made to apply the theory quantitatively to the swelling of proteins and other materials but it is particularly difficult to allow for the elastic properties of the gel itself (see also p. 233).

* PROCTOR, H. R., and WILSON, J. A., *J. Chem. Soc.* **109**, 307 (1916).

BIBLIOGRAPHY

ABRAMSON, H. A., *Electrokinetic Phenomena*, Chem. Cat. Co. (1934).

BOLAM, T. R., *The Donnan Equilibria*, G. Bell and Sons, Ltd., London (1932).

" The Electrical Double Layer." A General Discussion. *Trans. Faraday Soc.* **36**, 1–320 (1940).

FREUNDLICH, H., *Kapillarchemie*, Ed. 4, Akad. verlag. m.b.h., Leipzig (1930).

LOEB, J., *Proteins and the Theory of Colloidal Behavior*, McGraw-Hill (1924).

PAULI, W., and VALKÓ, E., *Kolloidchemie der Eiweisskörper*, Steinkopff, Leipzig (1933).

THOMAS, A. W., *Colloid Chemistry*, McGraw-Hill (1934).

J. Am. Leather Chemists Assoc. **29**, 3 (1934).

Chapter X

Gelation

A jelly such as that formed by cooling a solution of gelatin in water differs so fundamentally in properties from both ordinary liquids and solids that one must consider it as in a very real sense a distinct state of matter. The characteristics of jellies have long been a subject of study and dispute. Because of their industrial importance it is essential to have the clearest possible picture of jelly structure, but, to develop this, it seems advisable to call attention by way of introduction to the behavior of a series of related materials.

Plastic Flow.

The reader is familiar with ordinary mixed paint, consisting of a pigment suspended in a drying oil (p. 323). The paint will flow quite freely from the can, but, when properly applied to a vertical surface or even to the underside of a horizontal one, it remains in place, whereas an ordinary liquid, which would flow from the can even less readily, would drain off the surface to a considerable degree. Once movement is started, the paint flows quite readily, but when quiescent it offers definite resistance to initiation of flow. The phenomena have been described from the point of view of flow, but are equally important for the light thrown on the structure of suspensions. Figure 1 shows the torque *vs.* rate-of-shear relations of suspensions of ground glass in water, measured in a Stormer viscosimeter. Where the volumetric concentration of the suspended particles is below 34%, the suspensions show no abnormality of flow.* At about 37% a yield value begins to develop and an additional 2% of glass in the suspension increases this yield point to a value

---

* The viscosity is, of course, increased by the presence of the suspended particles.

higher than is necessary to maintain 500 R.P.M. in the instrument in the case of a 34% suspension.   In other words, a minimum volumetric concentration of suspended particles is essential for the exhibition of yield point, but, once this minimum is reached, relatively small increases in concentration induce large yield values.   The suspension has entered the region of plastic flow (p. 137).
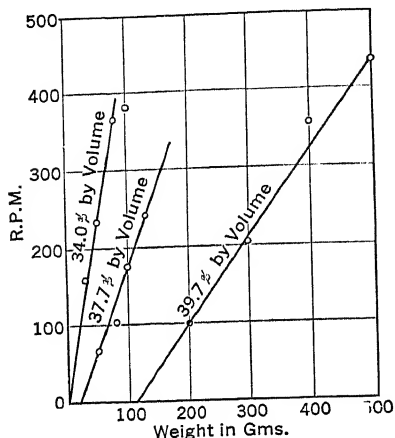
FIG. 1.   Weight-R.P.M. Curves for Suspensions of Ground Glass in Water. The open cup was used, and the experiment was conducted at room temperature.   In a 39% by volume suspension, the particles occupy their dry bulk volume.*

Granting the validity of the relations of Fig. 1, in the case of the 39.7% suspension a shearing force less than that corresponding to the 120 gms. yield value may cause distortion of the suspension, but will not induce flow; *i.e.*, the suspension acts as an elastic body, but not as a fluid.†   That in such a suspension the development of yield value may be due to the packing together of the particles sufficiently tightly so that they touch each other is indicated by the fact that not infrequently the lowest volumetric concentration at which a yield point appears corresponds to the concentration per unit volume under conditions of loose packing, as determined by pouring the dry particles carefully, without agitation, into an empty vessel.‡   Because there may be considerable concentration difference between conditions of tight and loose packing of the particles, there is a corresponding concentration range of plastic flow.§

* LEWIS, W. X., SQUIRES, L., and THOMPSON, W. I., *Trans. Am. Inst. Mining Met Eng.* **118**, 77 (1936).

† Abnormalities at shearing stresses in the neighborhood of the yield value and below are discussed elsewhere (p. 138).

‡ Indeed, attention may be called to the fact that a pile of sand has the flow characteristics of a suspension of this type.   It has a yield value corresponding to its angle of repose, and yet flows freely under higher shearing stresses.   Quicksand is a suspension of this type in which the voids between the sand grains are filled with water.

§ Where the particles are suspended in a liquid, the plastic range may be even wider, due perhaps to the lubricating effect of the liquid on the particles and to the possi-

In the development of yield value in mechanical suspensions the shape of the particles is of major importance. Thus, one can find samples of kieselguhr which are approximately 97% voids. If the voids be filled with liquid, the suspension has a yield point. The reason is that the exceedingly irregularly shaped particles of kieselguhr, consisting of thin, more or less film-like plates of irregular shape, interfere and touch each other at volumetric concentrations far below those corresponding to even loose packing of spheres. That irregularity in shape is frequently the cause of plastic flow in suspensions is indicated, for example, by the development of yield values in suspensions of the needle-shaped particles of vanadium pentoxide at re-markably low concentrations. Furthermore, the irregularities in shape need not be characteristic of the ultimate particles, but can be caused by irregular agglomerations. Thus, suspensions of exceedingly fine spherical glass particles of substantially uniform diameter in a liquid of the same density (to prevent sedimentation) have been found to exhibit yield values at con-centrations as low as 4% (cf. Fig. 2, p. 454), but whenever this occurs microscopic examination reveals extensive agglomeration of the spheres into clumps of irregular shape. The development of yield value at the low concentration is evidently due to the interference of the agglomerates.

Clay Suspensions.

That the behavior of the suspensions just mentioned is pri-marily mechanical seems beyond question, but there are other systems, remarkably similar in many respects, in which this conclusion is far less certain. Characteristic of these are suspen-sions of clays in water, particularly those made up of the clay particles of smaller size. Perhaps the most illuminating are the aqueous suspensions of the bentonites, significant properties of which are listed below:

1. The dry clay swells under water, in some cases to many times its volume, with imbibition of the liquid and substantially uniform distribution of it through the mass. There is consider-

bility of effective interference at concentrations even below that of loose packing of the dry particles, because of the increase in effective particle size caused by thin liquid films on the surfaces between the particles.

able affinity of the clay for the water, as indicated by the vapor pressure of water over it (Fig. 2) and by the swelling pressure exerted by the clay against a membrane permeable to the water but not to the clay itself. While this pressure drops off extremely rapidly as the clay imbibes water, the values at low water content may be many atmospheres.
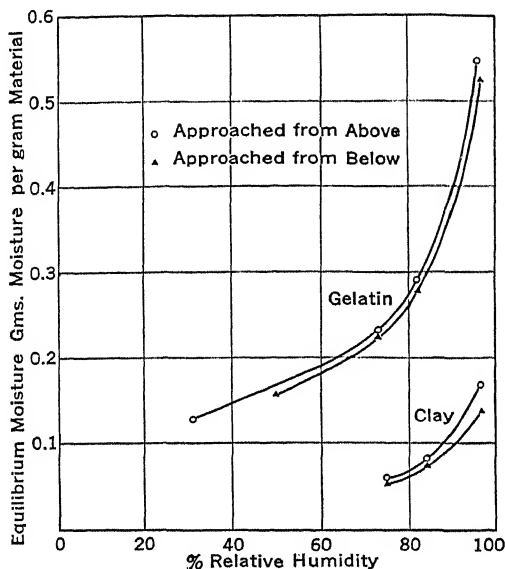
FIG. 2.   Vapor Pressure of Water over Clay at 25° C.*   (See p. 80.)

2. The clay suspensions, particularly c₄ the finer particles, can exhibit yield points and set to jellies or so-called gels at extremely low concentrations, even down to a tenth of a per cent of clay in water.* Furthermore, they can exhibit plastic flow over large ranges of clay concentrations, although if too little water be present under shearing stress the mass will disintegrate rather than flow.

3. The action of solvents is selective, *i.e.*, the character of the suspension is profoundly influenced by the nature of the liquid dispersing medium. Thus, a clay which disperses in

* HILL, N. C., *Mass. Inst. Tech. Thesis* (1928).
† While such gels are extremely weak mechanically, their existence can be demonstrated, as by the fact that a bubble of gas in the body of the suspension will be held there indefinitely, despite its tendency to rise.

water readily and to high degree will neither swell in oil nor disperse to anything like the degree in the aqueous suspension.

4. The clay sol is often profoundly sensitive to small quantities of electrolytes. Thus, the effect of specific addition agents on the flow characteristics of a bentonite purified by electrodialysis are seen in Figs. 12 and 13 on page 242.

5. While the mobility of a given clay sol drops with rise in temperature, the percentage rate of decrease is relatively low, often paralleling that of water (reminiscent of the Einstein equation, p. 131). Furthermore, the yield value of the suspension is usually relatively independent of the temperature.

6. Any given clay sol which is capable of setting to a gel has a setting time characteristic of the conditions.* However, the time required to set decreases as temperature rises, *i.e.*, rate of set increases with temperature.

7. Clay gels submitted to high shearing stresses break to smooth sols, with flow characteristics of the type of Fig. 15, Chapter VI, except at high concentrations, when they crumble and disintegrate.

Most clays do not exhibit all these characteristics in such high degree. Thus, the kaolins (see p. 450) swell less, exhibit plastic flow only at far higher concentrations, and are less affected by electrolytes. However, they too are selective in their action to various liquids and the effect of temperature on flow characteristics parallels that of the bentonites.

Gels.

An entirely different type of gels is formed by solution of such materials as gelatin, agar-agar, or the alkali soaps in water, or rubber in benzene. In appearance and mechanical properties they are often extraordinarily similar to the clay jellies. Like the clays, the dry materials swell spontaneously in the solvent to an extraordinary degree and exhibit high swelling pressure; they can give gel structures at extremely low concentrations and over wide concentration ranges; they are highly selective in

---

* This is the time necessary for the development of some arbitrarily chosen degree of mechanical rigidity, determined, for example, by noting the time just necessary to prevent flow on inversion in a tube of a given diameter containing the gel, or to prevent fall through it of a sphere of definite diameter and density in a vessel of given dimensions.

their action toward solvents,* and the sols are frequently sensitive to electrolytes. Despite these parallelisms in behavior, there are three fundamental differences. In the first place, up to relatively high concentrations, sols formed from one of these materials usually exhibit a temperature (or rather a narrow temperature range) above which the flow characteristics of the sol are normal, with no indication of yield point. Furthermore, at temperatures just below this range, flow measurements show yield values which increase and mobilities which decrease extremely rapidly with the temperature (Fig. 3). In the second
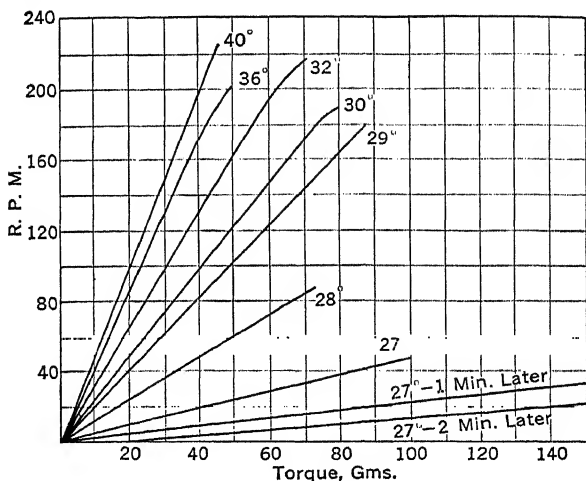
FIG. 3. Flow Relationships of a 20% Gelatin Solution Measured at Varying Temperatures (Stormer Viscosimeter).

place, while at high temperatures the sols do not set to gels, at low temperatures the rate of set increases extremely rapidly as temperature drops. In the third place, jellies of high strength, due either to low temperature or high concentration, tend to rupture and crumble, rather than to flow, when submitted to high shearing stress.

Before attempting the explanation of these complicated relationships, it seems better, in the light of the facts cited, first to offer a definition of a gel.

---

* Thus, gelatin, despite its affinity for water, is extraordinarily inert to most organic liquids. It is used to line alcohol barrels because of its impermeability to that solvent.

A gel is a mixture, one component of which is a fluid, homogeneous down to substantially colloidal dimensions, capable of resisting a finite shearing force.

Despite similarity to a gel in mechanical behavior, neither a sand pile nor a suspension such as kieselguhr in water comes under this definition, because the dispersed particles are not colloidal. A sponge or similar structure is excluded for the same reason, although, were its structure sufficiently fine, it would become a gel despite the fact that it constitutes a continuum rather than a dispersion of discrete particles. While many metals, such as copper and steel, show the flow characteristics of gels (except that the stresses required are enormously higher), they are not covered by the definition because their crystal grain structure is rarely, if ever, colloidal and there is no fluid constituent.* Glass is not included because the evidence indicates that its rigidity is due to extremely high viscosity rather than to a real yield value (see p. 299).

Gels are most frequently prepared by the gelation of a sol or by the swelling of the dry solute by imbibition of solvent. The fluid can be removed, as by evaporation if it be a volatile liquid. Where this is accompanied by a shrinkage corresponding approximately to the volume of liquid removed (as in drying gelatin gels), the product, though harder and stronger, often has the appearance of a gel, and is called a xerogel. Where the technique of liquid removal minimizes shrinkage but replaces the liquid by gas, the product is an aerogel (p. 227). A helpful classification of gels can be based on the type of sol from which the gel can be prepared. *Suspensoidal gels* are illustrated by the bentonitic clays, and *emulsoidal* gels by gelatin, agar-agar, rubber, and the like. The fundamental differences in the two types, already described, should be kept clearly in mind. Gels in which the sol-gel transformation can be repeated at will by suitable choice of conditions are called *reversible*, whereas those which when once set cannot be re-liquefied other than by indirect means are *irreversible*. The irreversible gels are most frequently of the emulsoidal type. Just as the distinction between suspensoids and emulsoids is indefinite, many sols showing over-

* Ignoring the possible influence of amorphous metal between the crystal grains.

lapping properties, so the classification of gels into suspensoidal and emulsoidal types, while helpful, must not be considered rigid. Furthermore, as will be illustrated in the following pages, a multitude of complicating factors are encountered which must be interpreted in the light of the facts of the individual case.

The swelling of dry, gel-forming materials in solvent usually evolves heat and is accompanied by a small but important shrinkage in total volume. Where the unimpeded swelling in excess solvent results in formation of a gel of a definite volume, the swelling is described as *limited*; where swelling proceeds to ultimate disintegration into a colloidal solution, it is *unlimited*. As already stated, if the swelling is opposed by a membrane (or its equivalent) permeable to solvent but not to solute, a swelling pressure develops at the surface of the membrane. Assuming reversibility of the swelling, the isothermal reversible work (free energy change) required for removal of dm units of mass of solvent from the swollen gel is $dW = Pdv$, where P is the swelling pressure and dv the change in volume of the gel. Neglecting the difference between dv and the volume of solvent removed $v_0\,dm$, where v_0 is the specific volume of the solvent, $dW = Pv_0dm$. The solvent may also be removed reversibly by evaporation at the vapor pressure p, of the solvent over the gel, followed by compression to saturation, and condensation. If the solvent vapors conform to the gas laws, this gives

$$dW = Pdv \quad \frac{RTdm}{M}\ln\frac{p_0}{p}.$$

Since the isothermal, reversible work must be independent of the path, the two quantities can be equated, yielding the relation between swelling pressure and vapor pressure of the gel,

$$P = \frac{RT}{Mv_0}\ln\frac{p_0}{p}.$$

A gel can be desolvated by freezing the solvent, even though this leaves the crystals formed dispersed through the mass. The freezing point can be determined by the usual technique of noting the point of temperature arrest in a properly conducted cooling curve, but unusual care in manipulation is necessary because of interference of the gel-forming material with crystallization of the solvent and the large change in partial pressure induced by separation of a small amount of crystals in gels of high concentration.

The analogy of swelling pressure to the osmotic pressure of sols is obvious. However, because the structure of the gel inhibits the complete freedom of movement of solute particles characteristic of the sol, it is not surprising to find considerable hysteresis effects (Fig. 2 and pp. 235 and 244). The affinity of the solute for the solvent, which is often very high indeed for the first increments of solvent imbibed (as indicated by swelling pressures that may amount to hundreds of atmospheres) but which falls off rapidly to

negligible values as swelling proceeds, is probably due far more to solvation resulting from partial valence forces than to the disgregating tendencies of thermal agitation, at least in the concentrated gels.

Properties of Gels.

Certain properties of gels throwing light upon their structure should be summarized.

1. That a gel can sustain a shearing stress without flow indicates that it must possess some sort of continuous mechanical network or structure. However, the phenomenon of gelation in excessively dilute solutions proves that this structure can consist of a relatively small fraction of the total mass,* a conclusion confirmed by the mechanical weakness of such gels.

2. Whenever it is possible to see Brownian movement of the particles in a sol prior to gelation, it is observed that when the sol sets to a gel the Brownian movement of the dispersed particles ceases. This indicates that the mechanical structure is made up of the colloidally dispersed particles of the sol, which become locked into the structure in the process of gelation.†

3. Experimental examination ‡ of dilute gels and the sols from which they are formed indicates substantial identity of all physical properties (such as vapor pressure, electrical conductivity, permeability to diffusing solutes, etc.), with the one exception of the mechanical rigidity of the gel in contrast to the fluidity of the sol.§ This apparently eliminates both phase inversion and the development of closed cells of the honeycomb type as mechanisms of gelation.

4. Kistler ‖ was able to replace the water of a gelatin gel by air without substantial collapse of the gel structure,¶ converting

---

* A wire bicycle wheel is a strong and rigid structure in which the structural members (*i.e.*, the wire spokes) occupy a very small fraction of the total volume. A gel must have some equivalent of such a structure, which, however, is three-dimensional and has little orientation of the members.

† Examined under the ultramicroscope, the visible particles of many dilute gels appear to be locked into position at considerable distances from each other. This does not necessarily prove lack of contact between the particles, because the shape and position of many of the particles relative to the illumination can easily make them invisible at a given angle of observation.

‡ LAING, M. E., and McBAIN, J. W., *J. Chem. Soc.* 117, 1506 (1920).

§ The tendency of the shearing stress *vs.* rate of shear curve to be continuous down to the yield value indicates a progressive transformation in structure from the relatively fluid sol at high rates of shear to the gel formed when shearing ceases.

‖ KISTLER, S. S., *J. Phys. Chem.* 36, 52 (1932).

¶ Kistler replaced the water of a dilute gelatin gel by progressive displacement with alcohol, replacing this similarly in succession by ether, petroleum ether, and propane

it into an aerogel.   This eliminates solvation, in the sense of combination, by adsorption or otherwise, of solvent with the gel-forming material, as a necessary condition for the continued existence of a gel structure once formed.

5. Some gels possess a considerable degree of "memory." Thus, assume gelatin dissolved in two solutions, a dilute and a concentrated one, both solutions being allowed to set as firm jellies without change in concentration, and the jellies then dried down at low temperatures to dry gelatin, so that they retain
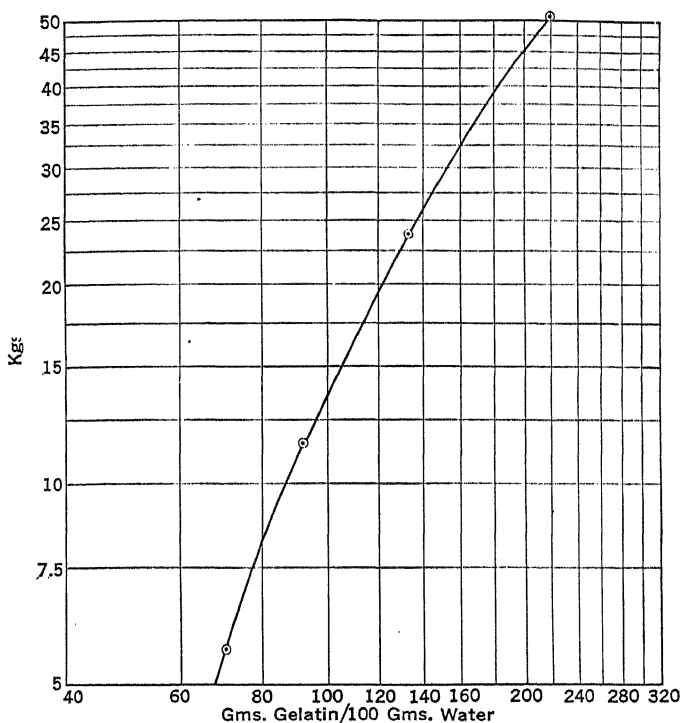
FIG. 4.   Swelling Pressure of Gelatin in Water.*

their gel structure throughout the drying operation.   On re-swelling in water at low temperature, the gelatin from the dilute jelly always tends to imbibe more water and swell further than

(at low temperatures), finally removing the last by heating under pressure to above the critical temperature, followed by release of pressure.

* Data from HILL, N. C., *Mass. Inst. Tech. Thesis* (1928),

the concentrated jelly.*   In many cases the gelatin will swell to almost the same volume as the jelly from which it was formed, although this depends to a high degree on the technique of manipulation and drying.   These facts indicate that the structure, even of the dried gelatin, is dictated in large degree by the conditions in the jelly at the time of set, emphasizing the stability of the structure developed in the process of gelation.

The Structure of Gels.

The only tenable explanation of the properties of gels in general seems to be the assumption of a card-house or brush-heap structure of solid or quasi-solid particles immersed in the fluid, the particles being bonded at points of contact sufficiently to prevent disintegration of the structure.   The particles of the frame work may be suspended solids of colloidal dimensions, or long, intertangled molecular chains, or, in limiting cases, even emulsified droplets.   The bonds at points of contact may be due in some cases to nothing more than interparticle friction, while in others the forces are undoubtedly of the partial valence or the primary valence types.   Despite the diversities in properties resulting from these possible differences in structure, it is felt that the picture here outlined gives the most satisfactory insight today available into the behavior of gels.   In the following, individual cases will be discussed in the light of this explanation.

In the case of gels built up of molecular chains, the parallelism with three-dimensional polymers (p. 155) is obvious.   Indeed, those polymers may be considered as xerogels of this particular type, formed, not by evaporation of solvent from the gel, but by direct synthesis from the components.   However, to secure gel structure in a brush-heap of molecular threads it is not necessary that they be held together by primary valence forces; far weaker forces of the partial valence type, or, conceivably, even the forces of interparticle friction, may suffice.   On the other hand, the nature of the bonding forces profoundly affects the strength of the gel (p. 244).

Behavior of Specific Gels.

Gelatin gels may be taken as typical of the emulsoidal type. Attention has already been called in Fig. 3 to the influence of

* Gelatin has a sufficient affinity for water so that when prepared by drying a concentrated jelly it will almost always swell in water to a volume greater than the jelly from which it was made.   In this connection compare Fig. 4.

temperature on their flow characteristics, and to their affinity for water as indicated by their vapor pressure and swelling pressure curves in Figs. 2 and 4.   The sol-gel transformation, corresponding to the development of yield values (see p. 220), is reversible except for a hysteresis effect of the order of magnitude of 5° C.   The affinity for water must be explained on the basis of some sort of solvation, *i.e.*, combination of water with the ultimate particle of protein.*   However, the mechanism of solvation is unknown; it may well consist of a layer of water molecules held by partial valence forces along and around the molecular chains of the gelatin, to the extent of perhaps two or three molecular diameters, equivalent to a kind of adsorption.   In any event, the water thus held greatly reduces the magnitude and effectiveness of the forces of mutual attraction of the gelatin molecules.   Moreover, while the affinity of the gelatin for the initial increments of imbibed water is very great, it drops off rapidly as the quantity of water held by the gelatin rises.   The reduction in the mutual attractions of the gelatin molecules parallels this change.
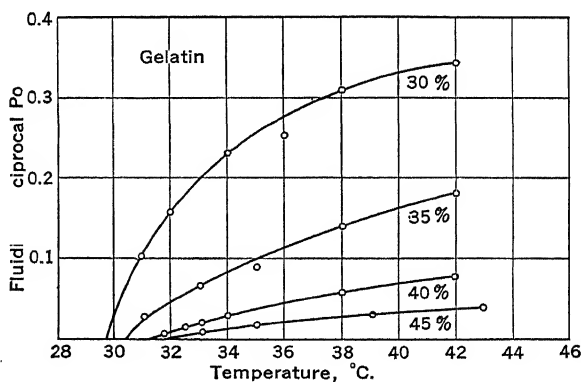
FIG. 5.   Fluidity-Temperature Curves for Gelatin.†

The normal flow characteristics observed at high temperatures (Fig. 5) show that under these conditions in sols of the concentrations in question the mutual attractions of the gelatin

---

* That the imbibition of water cannot be a matter of capillarity is evidenced by the fact that only highly polar liquids affect the gelatin, ordinary organic liquids being remarkably inert toward it (see p. 233).

† WIEDEMAN, O., *Mass. Inst. Tech. Thesis* (1935).

molecules are negligible, but the incidence of gelation in the neighborhood of 30° C. apparently makes it necessary to assume that, at these temperatures and below, the affinities are sufficient to cause molecular adhesion, at least at many points of contact. This is readily explained in the light of the fields of force which must exist around the gelatin molecule. The partial valence forces cause attractions between the particles which, however, particularly in the presence of the water layer surrounding the particles in an aqueous sol, are not large. While these forces tend to hold the particles together wherever they happen to come into contact, there is the simultaneous disgregating effect of thermal agitation, as evidenced in Brownian movement. In the case of gelatin at temperature levels above about 30° C., the thermal forces preponderate, and any adhesion which may occur on contacts of gelatin molecules lacks permanence. As temperature drops, the tendency toward thermal disgregation decreases, and, since intermolecular attraction remains constant or may even increase slightly, a temperature is reached at which intermolecular contacts once established tend to remain permanent. In the case of a colloidal solution of a more or less linear molecule of high molecular weight, such as gelatin is assumed to be, this produces an interlocking network of molecular fibrillae, sufficient to give considerable mechanical rigidity to the mass, but leaving such a high percentage of open spaces filled with liquid that such properties as electrical conductivity, diffusivity of solutes, and the like, are essentially unaffected. This explanation makes clear the reason for the fact that the structural strength of a gelatin gel, as evidenced by the yield value of the sol from which it is formed, increases rapidly as temperature falls. The disgregating tendencies of thermal agitation are reduced, so that the net attractions binding together the molecular fibrillae increase. Figure 5, based on flow measurements all taken above gelation temperature where no yield point is detectable, indicates, as is to be expected, that these molecular associations do not cease abruptly at the gel point but persist at higher temperatures, causing the rapid change in fluidity with temperature seen in the figure.* In the case of agar (Fig. 6), the effect is

---

* This is in marked contrast to the behavior of clay gels (p. 223). The method of plotting here used offers a good method of estimating temperature of gelation.

still more striking.   The explanation likewise makes clear the
effect of concentration on temperature of gelation (Fig. 7).   At
temperature levels at which the tendency toward gelation of the
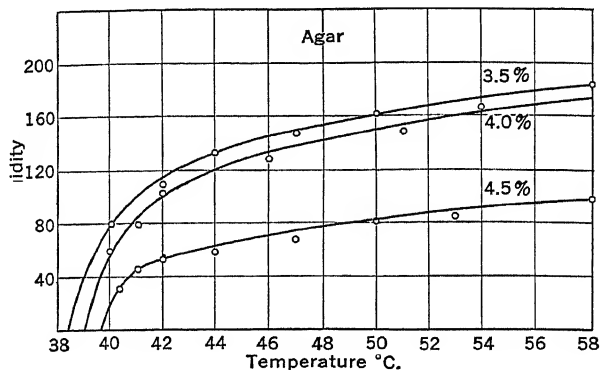
FIG. 6.   Fluidity-Temperature Curves for Agar.*

sol is slight (say, a little above 30° C.), only a small fraction of
the molecular contacts that occur is permanent enough to be
effective in developing structure.   In a concentrated solution
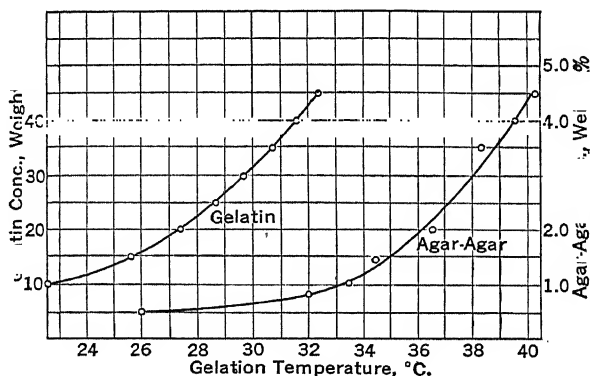
FIG. 7.   Concentration-Gel Temperature Curves.*

there will still be enough effective contacts to interfere with flow
and indicate a yield point, but in dilute solutions this will not
be true.   It will be noted from Fig. 7 that the gelatin there
examined showed a gelation temperature range of approximately
10° C., depending on the concentration.

* WIEDEMAN, O., *Mass. Inst. Tech. Thesis* (1935).

The effect of non-electrolytes miscible with water is instructive. Alcohol poured over a gelatin jelly will shrink it and, if sufficient in amount and concentration, dehydrate it almost completely.* On the other hand, alcohol added to a gelatin sol precipitates it, flocking it out in curds or strings. This is certainly also a dehydrating effect on the ultimate gelatin molecule, offering further proof that solvation of the particles by water is an essential element in dispersion in the aqueous sol.

FIG. 8. Cowhide in Dilute Acid.† Basis: 1 gram Hide Substance. (Original volume of neutral hide = 3.77.)

Some of the most important characteristics of these gels depend upon the amphoteric properties of the gelatin. The gel reacts with acids and bases, the distribution between these and any phase external to the jelly itself being dictated by the Donnan equilibrium (p. 212). Figure 8 shows the amount of monovalent acid absorbed by cowhide ‡ as a function of pH and the corresponding volume or swelling effect. Figure 9 shows the equivalent data for a gelatin gel in HCl, based on measurements of Proctor and Wilson. ‖ The parallelisms of Figs. 8 and 9 are obvious, but the differences significant. The gelatin swells over

FIG. 9. Gelatin in Dilute Acid.§ Basis: 1 gram Dry Gelatin.

* If a very dilute solution of alcohol be brought in contact with the gel, the solute diffuses through it and distributes itself quite uniformly. Moreover, by treating with progressively increasing concentrations of alcohol, keeping concentration difference very low, the alcohol will replace the water completely, without much shrinking of the gel. (Note the work of Kistler, p. 227.)

† STUCKLEN, H. W., *Mass. Inst. Tech. Thesis* (1911).

‡ Analogous to gelatin (p. 211), though swelling less because of its strength.

§ PROCTOR, H. R., and WILSON, J. A., *J. Chem. Soc.* 109, 307 (1916).

‖ They do not give their data for high pH values, other than in graphical form difficult to evaluate; the corresponding range in Fig. 9 is dotted.

seven times as much as hide and, largely on this account, imbibes far more acid at a given pH. In the case of a gel of this sort, the gel structure itself acts as its own semipermeable membrane. The excess of diffusible ions taken up by the gel, as dictated by the Donnan equilibrium, generates an osmotic pressure, the tendency of which is to expand the gel. However, the gel itself has a tensile strength and a corresponding resistance to increase in volume. Proctor and Wilson (p. 217) have studied the relationships quantitatively, calculating the effective osmotic pressure from the Donnan equilibrium on the assumption that the gelatin behaves as a monovalent electrolyte, and assuming that this osmotic pressure is counterbalanced by a resistance to expansion of the gelatin gel which is proportional to its increase in volume.* They have thus been able to show substantial correlation between the observed swelling and the acid imbibition of gelatin.

An important distinction between the behavior of an ampho-teric gel in equilibrium with an external phase and that of the corresponding sol under osmotic equilibrium against a semi-permeable membrane should be emphasized. The osmotic pressure in the sol is high but uniform. The pressure differential between the sol and the external liquid develops at the membrane substantially discontinuously, and this osmotic pressure must be counterbalanced by mechanical stresses in the membrane itself. In the gel in equilibrium with its surrounding liquid there also exists an excess osmotic pressure. However, this is not counterbalanced by a stress in the interface between the gel and the liquid around it and there is no sudden drop in pressure at this surface. The expansive tendency of the osmotic pressure is counterbalanced by stresses in the framework of the gel itself, *i.e.*, in the case of gelatin, by tension in the molecular fibrillae. Considering any differential volume in the gel, which is large relative to the size of an ultimate gelatin particle, the stresses are counterbalanced completely within each of these units of volume, a situation entirely different from that existing in a sol exerting its osmotic pressure against a semipermeable membrane.

---

* The argument in favor of this assumption is not entirely convincing. Study of Fig. 9 shows that it is incompatible with the data of that plot.

The so-called Schroeder paradox is the fact that a gelatin xerogel will swell more in liquid water than in contact with saturated water vapor at the same temperature. Granting that the gel swollen in vapor can still exert a finite swelling pressure in liquid water, the phenomenon would constitute a contradiction of the second law of thermodynamics, since it could be used as a mechanism for isothermal conversion of heat into work. The facts are in dispute because satisfactory experimental manipulation is difficult.

Agar-agar is an emulsoid from seaweed, of complicated molecular structure containing sulfonic radicals and carbohydrate residues. It is in many respects similar to gelatin, but has far higher jelly strength at a given concentration, has a higher temperature range of gelation (Fig. 7), and is less sensitive to changes in pH. It exhibits high hysteresis in gelation, in that to liquefy its gel it must be heated well above the temperature of gelation.

Fatty Acid Soaps.

The alkali salts of certain of the straight-chain fatty acids form interesting gels of great industrial importance. Salts of acids of short chain length are water-soluble and behave as normal crystalloidal electrolytes. Only in the case of chains of about 8 carbon atoms and above do the abnormalities characteristic of the ordinary soaps appear, increasing as chain length goes up. Sodium palmitate may perhaps be considered as typical. At low temperatures (0° C.) it is relatively insoluble, but is very soluble at 100° C. In dilute aqueous solutions its behavior is normal; vapor pressure lowering and electrical conductivity indicate little or no molecular association of the salt and considerable electrolytic dissociation (although somewhat low for salts of this type). At high concentrations both vapor pressure lowering and conductivity are abnormally low; indeed, there is a narrow range in which vapor pressure lowering actually decreases with increasing concentration. Evidently, in concentrated solutions the soap molecules associate to high degree to micelles of colloidal size. Some at least of these micelles are charged, but they apparently consist to a considerable degree of neutral soap. Above about 70° C. these colloidal solutions are stable, but as temperature falls below this level the solutions

tend to deposit progressively increasing amounts of a so-called curd,— long fibrillae consisting, predominantly at least, of highly hydrated neutral soap. As the temperature is lowered a concentrated solution can transform itself into an opaque mass of curds, the interstices between the particles of which, however, will be found filled with residual soap solution or gel. At a temperature of 0° C. the soap will ultimately exist almost exclusively in the form of curd, its concentration in the external liquid being very small. Within these temperature ranges the solutions can under proper conditions be obtained in the form of gels, nearly transparent and homogeneous, in contrast to the relatively opaque curds.

The mechanism of molecular association which converts concentrated soap solutions so predominantly into an emulsoid colloid is uncertain, but the capacity of soaps to function as detergents is undoubtedly intimately associated with it (p. 261). In general, the higher the molecular weight of the soap, the lower its solubility at a given temperature level; increasing the unsaturation increases solubility for a given chain length. Furthermore, many other materials, such as the rosin soaps, many sulfonic acids of high molecular weight, and alkyl sulfates of long chain length, exhibit similar behavior (see pages 264–265).

Gels in Organic Solvents.

Rubber cements, solutions of rubber in suitable organic solvents, offer an illustration of a characteristic non-aqueous gel. Their properties are profoundly influenced by the rubber they contain, one of the most important factors being chain length (p. 143), as indicated by the viscosity of the solution (p. 165). The behavior of a typical solution in the neighborhood of its temperature of gelation is shown in Figs. 10 and 11.* While above a temperature of about − 40° C. the sols show normal characteristics of flow, as this temperature is approached there is an abnormally rapid fall in fluidity, indicating a zero value at about − 41° C. In fact, below this temperature, flow becomes anomalous, a yield value being indicated by the shear values at high rates of flow; moreover, on allowing the solution

---

* Note the parallelism to Figs. 3 and 5.

to stand, incipient gelation develops. In other words, rubber sols, like gelatin, have a gelation temperature or temperature
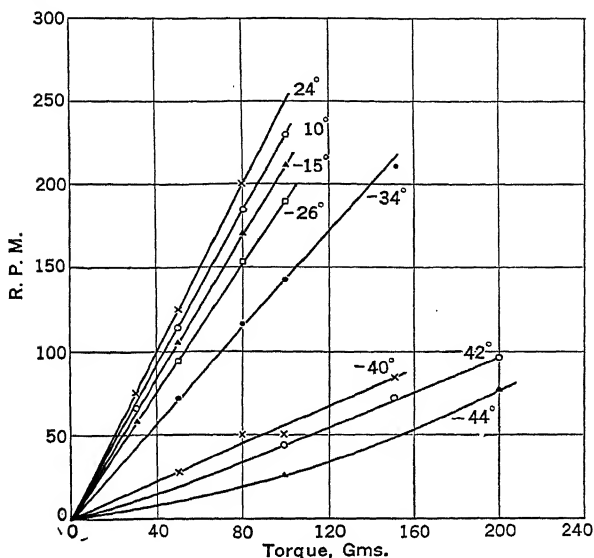
FIG. 10. 3% Rubber in Carbon Disulfide.*

range above which gelation does not develop. This temperature is greatly affected by breakdown (p. 413) of the rubber; for high-grade crude rubbers it is well above normal temperature, but is easily lowered, e.g., by milling the rubber (p. 411).

Rubber gels illustrate the specific effects of solvents. Thus, while gelatin can be dispersed and converted into a dilute gel only by the use of highly polar solvents such as water,† rubber responds only to solvents of relatively low
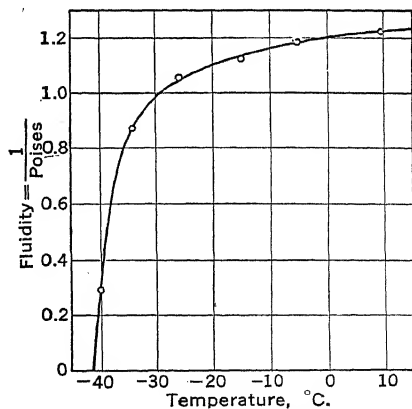
FIG. 11. 3% Rubber in Carbon Disulfide.*

* CHISWELL, E. B., JR., *Mass. Inst. Tech. Thesis* (1935).
† Glycerin is an effective solvent for gelatin and is used as such in hectograph compositions and printing rolls.

polarity, such as aromatic hydrocarbons,* carbon disulfide, etc. Moreover, just as gelatin can be precipitated or its gels desolvated by solvents of low polarity, miscible with water, so rubber gels can be desolvated by solvents of relatively high polarity, such as acetone and alcohol (see p. 15).

Irreversible Gels.

The parallelism between gels of the molecular chain type and three-dimensional polymers has already been mentioned. In any such gel, in which the bonds between the chains are of the primary valence type, one would expect to have extreme stability. Thus, conversion of the gel into a sol by mere dilution, heat, or agitation should be impossible. Even though the gel were originally formed from a sol, it can no longer be reconverted to it, *i.e.*, the gel has become *irreversible*. Thus, the typical reversible gelatin gel can be converted into an irreversible form by treatment with formaldehyde, undoubtedly through interaction with basic nitrogen groups to form stable molecular bridges between the chains. The most important irreversible gels are inorganic in structure; of them, silica gel may be taken as typical.

Silica gel is formed by mixing strong solutions of sodium silicate and a mineral acid, *e.g.*, hydrochloric.† The immediate product of the reaction is an emulsoidal sol of hydrated silicic acid dispersed in the solution of neutral salt. If concentration is sufficiently high, on standing quietly for a long period ‡ the sol transforms itself progressively into a gel of high mechanical strength and rigidity. After gelation is complete the salt can be washed out, leaving a solid of low density containing large amounts of water removable by drying, which yields a strong, highly porous gel.

---

* Paraffinic hydrocarbons are poor solvents for rubber. Their polarity is apparently too low.

† A good technique of mixing is to run equivalent streams of silicate and acid slowly but continuously into a large vessel which is equipped with an efficient stirrer. and full of reaction product, overflowing continuously into suitable receivers. The liquid in the vessel is completely neutralized, but can serve as diluent for the reactants, so that these interact with each other under conditions of extremely high dilution, thereby eliminating the coagulating effects of concentrated solutions of the reactants, both upon each other and upon the reaction product.

‡ If the material be agitated in the early stages of setting, the gel is disintegrated and never again regains its full original gelation capacity.

The silica undergoes progressive changes, not only during the process of setting, as indicated by development of mechanical strength, but also during drying, as evidenced by a shrinkage which, though relatively small, is progressive. Moreover, in the earlier stages of drying, moisture once lost by evaporation cannot be restored at the same conditions of temperature and partial pressure of water vapor, *i.e.*, there is a hysteresis effect in water removal in this stage of drying. The product dried under normal conditions is still by no means pure silica, but retains water removable only at elevated temperatures. The gel possesses high adsorptive capacity, particularly for water and other polar vapors, evidently due to its low density and porous structure (see p. 75).

The mechanism of gelation is not certain, but is doubtless due to the development of primary valence oxygen bridges between silica atoms, as discussed on pp. 283 to 286. It is clear that the residual hydroxyls offer the possibility of unlimited further condensation, not merely of the linear type, but three-dimensional as well. It is thus not surprising to find that silica gel possesses many of the properties of characteristic three-dimensional polymers. Keeping in mind the dilution of the silica in the aqueous menstruum at the time of gelation, it is obvious that the silica residues are locked into position relatively permanently by the mechanism described, affording an explanation of the "memory" of the gel, in the sense that it permanently retains in large degree the shape in which it is originally synthesized.

The technique of preparation and the properties of silica gel depend on the fact that, while the silicate ion in aqueous solutions of the alkali silicates is highly hydrated and is liberated by the action of strong acids as some form of hydrosilicic acid, this acid is inherently unstable. Although its tendency is to dehydrate even in the presence of excess water in dilute solution, in neutral solutions at normal temperatures the rate of dehydration is very low. Furthermore, as is true of so many weak, inorganic oxyacids, dehydration tends to occur preferentially between adjacent molecules rather than within a single molecular unit. However, as dehydration progresses, further dehydra-

tion rate becomes excessively slow and may practically cease altogether.*

Other irreversible gels similar to silica gel in methods of formation and properties and presumably also in structure are prepared from various heavy metal hydroxides, such as alumina and ferric hydroxide.

Suspensoid Gels.

The clay gels may be taken as typical of the suspensoid type. In addition to the characteristics already described (p. 221), emphasis should be laid on their electrical properties and mechanical behavior, the latter affording an illustration of an important property of matter.

While the clays are hydrated silicates of alumina, they are not chemically neutral but amphoteric, acid properties far outweighing basic tendencies. Natural clays contain metals, particularly alkalies and alkaline earths, present predominantly as salts of the clay functioning as an acid; clays of the kaolinitic type hold relatively small amounts, whereas bentonites may contain as much as 10%. The combined water of the clay is evidently present as hydroxyl groups, the hydrogens of which are acidic and correspondingly replaceable. The combined metals can be removed, largely at least, by acid washing or electrodialysis of the aqueous suspension of the clay. The initial stages of washing of an acid-treated clay offer little difficulty, because the acid flocculates the clay, but as the acid concentration falls with progressive washing, the clay, particularly if of a plastic, highly colloidal type, disperses so that further washing by either sedimentation or filtration becomes extremely difficult. Electrodialysis is best conducted in a cell with double diaphragms, between which the suspension of clay is placed. The metals are carried by the current to the cathode compartment and any soluble anions to the anode, whence they are removed by flowing streams of distilled water through these compartments. The clay itself, owing to its negative charge caused by its acidic ionization, travels towards the anode but is restrained by the

---

* This may be due to the fact that finally the residual hydroxyls are located at points in the structure too far apart for reaction to occur.

anodic diaphragm.   Best results are usually achieved by keeping the clay between the diaphragms in suspension by suitable agitation.   No fully satisfactory technique of industrial electro-dialysis has been developed.   The operation is slow at best. Because of the very great increase in resistance as the cations are removed, it is almost essential to operate at progressively increasing voltages.   The resistance finally becomes so high that the voltage must be raised to excessive values (often 100 to 200 volts), involving high power consumption per unit of metal removed; else the current falls so low that purification rate is inordinately slow.

Purification is never satisfactory, due in the case of acid treatment to the difficulties of washing and in electrodialysis to the low solubility of the metal hydroxides, especially magnesia.* The properties of clay dispersions are profoundly influenced by the combined metals.   Thus, the effect of barium and sodium hydroxides on the yield value of a suspension of bentonite purified by electrodialysis is shown in Fig. 12 and the correspond-ing curves for a kaolin in Fig. 13.   The latter figure also shows the influence of the barium hydroxide on the mobility.   The ultimate cause of the effects is unknown, but no explanation of the behavior of clay suspensions can be satisfactory unless it accounts for them.

X-ray examination of the clays has led to the conclusion that the ultimate particles are crystallites of the probable structure shown in Fig. 16.   It will be noted that this structure accords well with the chemical and physical behavior just described.

The behavior of clays throws important light on their struc-ture.   Natural bentonites, containing as they do large quantities of combined alkalies and alkaline earths, disperse readily and reversibly in water.   If, however, such a dispersion, particularly of the fine particles, be purified by electrodialysis, reversibility is lost, although the degree of dispersion is not markedly affected. When the metal-free dispersion is dried down, the dry "clay acid" will no longer redisperse, even on grinding, to anything

* It has been suggested that some of the combined metals, particularly the magnesia, are held within the ultimate particles of clay, through which they can diffuse either not at all or only with difficulty, the removable metals being combined on the surface of the particle.

like the degree characteristic of the original clay.   Furthermore,
if a drop of a dilute suspension of the finest particles be put on
a microscope slide under a cover glass and allowed to dry out,
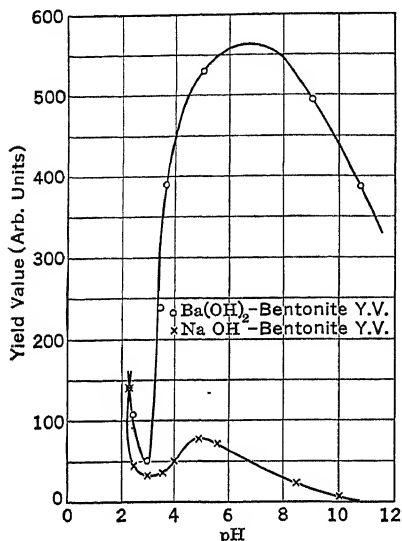
FIG. 12.   Viscous Characteristics
of Suspensions.*
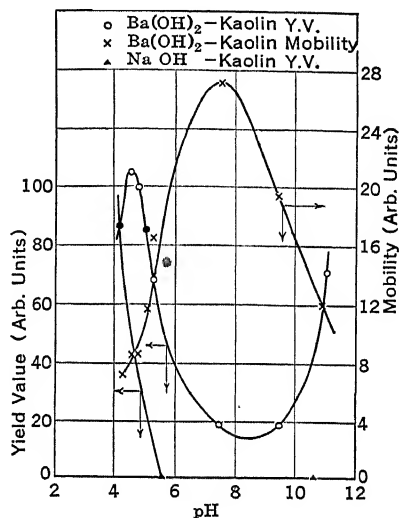
FIG. 13.   Variation in Yield and
Mobility with pH.*

it is found that if one focuses on the receding edge of the evapo-
rating water film, while originally no particles are visible, platy
particles begin to develop in the evaporating surface, as though
crystallization were occurring at this point.   Evidently, as the
water recedes, platy particles coming into the surface are oriented
parallel to it, and in this way build up a pack of cards which
ultimately becomes visible and develops into a grain of dry clay.
Apparently, if the platy particle of clay is coated with ions of
combined metal, two plates coalescing face to face are separated
sufficiently so that water can penetrate between them and re-
disperse them.   If, however, the particles are coated only with
hydrogen ions, the volume of these is so small that on drying
two plates face to face they approach too closely for water
penetration to occur.

A clay carefully purified by electrodialysis drops in pH to a value characteristic of the sample. Thus, a bentonite with a pH in its natural state of 9 or even 10, on electrodialysis, can drop as low as 2.2, the exact value depending on the concentration. However, it is found that the conductivity of this clay suspension is only a small fraction, 20 or 25%, of that corresponding to the mobility of a hydrogen-ion solution of concentration equivalent to the pH. The hydrogen ion in the electrodialyzed clay, while capable of affecting a hydrogen electrode, is for some unknown reason anchored to the clay particle in such a way that it is incapable of carrying current to the normal degree.

Thixotropy.

Bentonite gels, particularly the very dilute ones, often liquefy readily on agitation to a fluid suspension of relatively low viscosity, the gel reforming reversibly on standing. The phenomenon is called thixotropy and is characteristic of many gels, but not of others. Typical emulsoidal gels usually exhibit it to only a slight degree, tending to disintegrate rather than liquefy under shear, particularly if the gels are mechanically strong. The property is of industrial importance, e.g., in paints (p. 327).

An insight into the phenomenon is gained by study of Fig. 14. The three upper curves of that figure are obtained by pouring a thoroughly mixed 10.7% aqueous suspension of ball clay into a MacMichael viscosimeter, allowing it to stand for the variable time indicated on each curve, then starting up the instrument at a constant speed of 105 R.P.M. and reading the torque on the rotating cylinder as a function of the time of stirring. The time of standing before initiation of rotation has but little effect on the reading of the instrument. Indeed, the maximum deviation between the three curves is only 6%, probably little if any outside the reproducibility of the measurements. However, the viscosity of the suspension (proportional to the torque) rises about 60% after stirring starts. The rate of rise is at first rapid but falls off as the ultimate value is approached, about fifteen minutes being required to reach a substantially constant value. Clearly, the fluid suspension has a more highly developed structure (as measured by resistance to flow) after stirring than

before.* In other words, the suspension has a structure the nature of which ultimately depends upon the rate of shear (see Fig. 1), but if the suspension, submitted to a definite rate of shear, starts out with a structure different from this ultimate
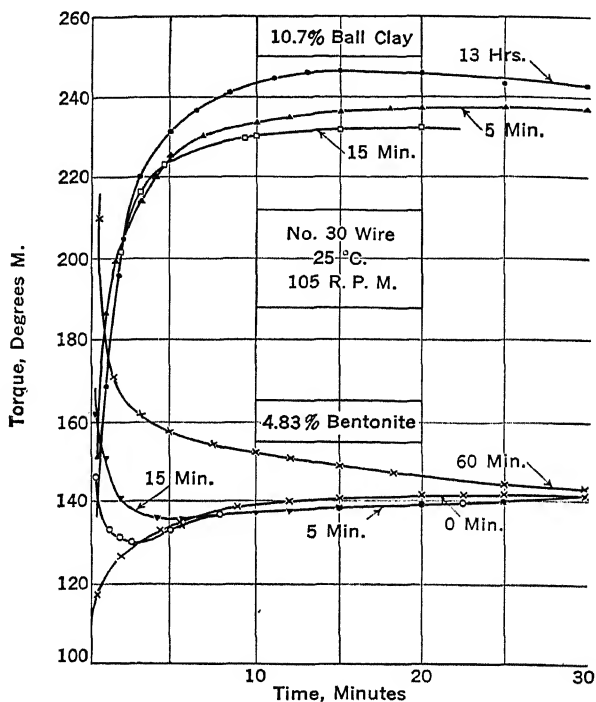
FIG. 14. Effect of Agitation on Viscosity of Bentonite Gels.†

value, the structure transforms itself relatively slowly to the final equilibrium condition. However, this transformation does not occur on standing; it evidently requires the movement or agitation involved in flow to enable it to proceed.

The importance of this behavior in the determination of the flow characteristics of suspensions is obvious. If the material happens to be one in which the curve just discussed rises to the asymptotic value extremely

* On the other hand, the relatively low initial viscosity is undoubtedly due to the agitation involved in pouring the suspension into the instrument, which is even more violent in its disruptive effect than that corresponding to the rate of rotation employed in the measurements.

† BROUGHTON, G., and HAND, R. S., Am. Inst. Mining Met. Eng., Tech. Publication No. 1002, and Petroleum Technology, Nov., 1938.

rapidly no difficulty is encountered.   If, however, as in this case, there is a time lag, *i.e.*, a hysteresis effect, the measured torque will depend not only on the time of stirring before the measurement is taken but also on the history of the suspension prior to introduction into the instrument.   In such event it is usually best to determine the asymptotic viscosity as a measure of the ultimate equilibrium structure of the suspension.   The unsuitability of an instrument of the Stormer type (p. 22) is obvious.

The behavior of a characteristic bentonite suspension (in this case 4.83% clay) contrasts sharply with that of the ball clay.   Here, too, the viscosity (torque) comes to an asymptotic value independent of the time of stirring.   However, whereas the curve for zero time of standing has the same general shape as the curve for the ball clay suspension, initial standing of the bentonite completely changes the curve.   The initial torque rises sharply on standing; even five minutes is enough to raise the initial torque to a value higher than the ultimate asymptote.   Sufficient standing time (sixty minutes) reverses the whole curve in the sense that the torque

FIG. 15.   Effect of Quebracho Tannin on Yield Point of a Clay Mud at Various pH Values.*

now drops progressively with stirring time.   Intermediate standing times give an initial drop in viscosity with stirring, followed by a minimum value and an ultimate rise to the equilibrium asymptote.

The bentonite suspensions here tested possess yield values, *i.e.*, on standing they set to gels.   Inspection of the curves makes it obvious that the viscosity of the suspension formed by long standing drops exceedingly rapidly with time of stirring, even at constant stirring speed.   In other words, stirring makes the suspension much more fluid, but if the stirring be stopped the viscosity again rises and gelation occurs, the whole process being reversible.   This phenomenon constitutes thixotropy.

* Lewis, W. K., Squires, L., and Thompson, W. I., *Trans. Am. Inst. Mining Met. Eng.* 114, 42 (1935).

The significance of the minima observed in the curves for intermediate times of set requires emphasis. If standing merely allowed the suspension to develop the structure characteristic of the asymptotic condition achieved after long continued stirring, while it might well develop to a far higher degree than in the moving solution, the movement of stirring at constant rate of shear would merely break the structure down to the equilibrium asymptotic condition; it could not possibly break it down below the asymptotic value and give a minimum in the time curve, followed by an ultimate rise. In other words, the minima force the conclusion that these suspensions possess at least two different types of structure. The first type, similar to that exhibited by ball clay, can develop by stirring. The second type, absent in ball clay, can develop on standing in the quiescent suspension. However, in these cases the second type is broken down very rapidly by stirring, relative to the rate of build-up of the first. If, therefore, sufficient time of standing has not been allowed for development of excessive structure of the second type, a short time of stirring will break this down sufficiently before the first type of structure has had a chance to develop, so that the viscosity of the whole mass is lower than the asymptotic value, *i.e.*, the time curve exhibits a minimum.

Measurements similar to those of Fig. 14 made at various rates of shear show that the asymptotic torques plotted against rate of shear correspond to plastic flow with a high yield value. In other words, if the first type of structure be developed by stirring, and the rate of stirring slowly reduced, one will obtain a gel requiring finite shear to initiate flow, entirely aside from the development of any structure of the second type.

Assuming constant rate of rotation of the viscosimeter, calling the time of rotation θ, the corresponding torque T, and the asymptotic torque T_∞, it is found that the torque can be expressed by the equation,

$$T = T_\infty + T_0'e^{-k\theta} + (T_0 - T_\infty)e^{-K\theta}.$$

T_0 and T_0' are constants, the former evidently equivalent to the initial value of the shear required to overcome the first type of structure and T_0' to the second. For the clay in question k is far larger than K, indicating rapid breakdown of thixotropic structure. The data prove that both T_0 and T_0' increase with time of standing. The difference between T_0' and the value of T_0 for zero time of set is the magnitude of thixotropic set; k can be considered the coefficient of thixotropic breakdown. The breakdown of struc-

ture is evidently a first order reaction, in the sense that the rate of breakdown is proportional to the distance from equilibrium.

In certain thixotropic sols, including bentonites, rate of set on standing is greatly increased by vibration (*e.g.*, tapping), a phenomenon known as rheopexy.

The nature of the structure of clay gels is a matter of dispute. It is difficult to visualize a mechanism of gelation which can occur in excessively dilute bentonite suspensions (p. 222). Because the development of yield point and gelation in very dilute
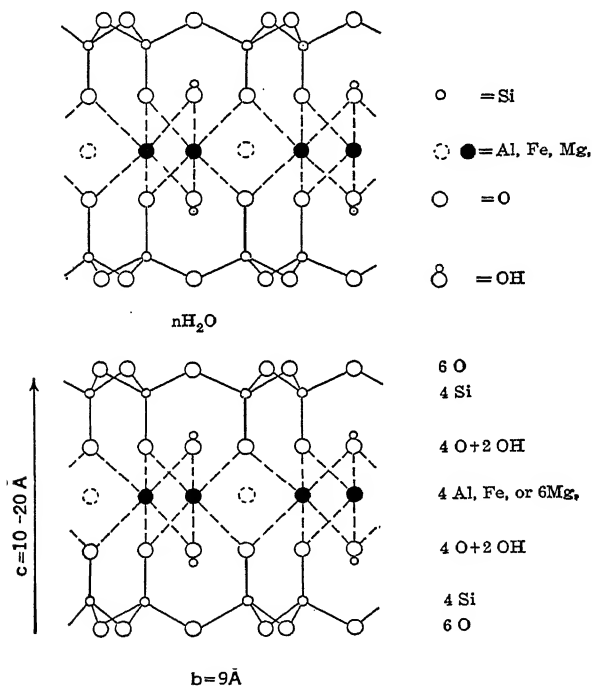
FIG. 16. Probable Structure of Ultimate Clay Particle.* (Montmorillonite.)

suspensoidal dispersions never occurs unless the particles are highly anisometric, it is quite generally agreed that the gel structure is at least in some degree augmented by irregular particle shape (pp. 221 and 454). In the case of the platy clay particles, one can visualize a structure consisting of interfering cards, locked into place by friction at the points of edge-to-edge

* Hofmann, U., and Bilke, W., "Die Quellung und das Basenaustauschvermögen des Montmorillonits," *Kolloid Z.* 77, 238 (1936). (Theodor Steinkopff.)

contact. If, however, this is the only mechanism of develop-
ment of structure in these platy suspensions, to explain gelation
at a concentration of 0.1% or even less makes it necessary to
assume plates in which the ratio of length and width to thickness
is several hundred. In view of even the minimum thickness of
the clay particles indicated by Fig. 16, such extreme ratios of
dimensions seem highly improbable, and all data at present
available confirm this point. However, attention has already
been called to the possibility of developing structure in relatively
dilute suspensions of spherical particles (p. 454) by agglomera-
tion, and it seems highly probable that the capacity of the platy
clay particles to build up structure can be greatly enhanced by
this factor. Confirmation is found in the fact that the addition
in small amounts of water-soluble materials, particularly electro-
lytes, which in large concentration cause a high degree of floccu-
lation, often induces gelation.* It seems probable that the
gelation of clay suspensions is due to the mechanical interference
of the platy particles, which may be greatly enhanced by particle
agglomeration, with the stability of the structure depending
largely on the amount of interparticle friction at points of con-
tact. Agglomeration and friction are in turn profoundly influ-
enced by the metals combined on the particle surface and by the
electrolytes in the solution, leading to extreme variations in the
relative effects of the different factors and consequent complica-
tions in behavior of the gels.

The yield values and gelation tendencies of clay suspensions are pro-
foundly influenced, not only by pH and the specific alkali combined with
the clay (Fig. 12), but also by certain anions in the solution, usually of high
molecular weight. Thus, addition of tannic acids to clay suspensions
alkaline in reaction can affect the yield values to an extraordinary degree
(Fig. 15); advantage has long been taken of this effect to secure fluid clay
suspensions (slips) at high concentrations for use in ceramic manufacture
(p. 456). Soluble hexametaphosphates often exhibit somewhat similar
behavior. The quantities required are often small relative to the amount
of clay treated and their use is common wherever control of the flow charac-

---

* Thus, extremely dilute suspensions of electrodialyzed, relatively monodisperse,
colloidal bentonite, which exhibit little or no gelling tendency, can be converted into firm
gels by the addition of small amounts of KOH or KCl. The necessary concentrations of
the added electrolytes are low, but their ratio to that of the clay is high, from 1 to 5 mols
of electrolyte per equivalent of clay (equivalent weight approximately 1100). The same
electrolytes added in high concentration to the initial suspensions flock out the clay.

teristics of clay suspensions is desirable, as in the muds employed in the drilling of deep wells.

BIBLIOGRAPHY

ACADEMY OF SCIENCES AT AMSTERDAM, *First and Second Reports on Viscosity and Plasticity*, Amsterdam (1935 and 1938).
ALEXANDER, J., *Colloid Chemistry*, Vol. I, Chem. Cat. Co. (1926).
HATSCHEK, E., *Introduction to Physics and Chemistry of Colloids*, Ed. 5, Churchill, Ltd., London (1925).
HEDGES, E. S., *Colloids*, Edward Arnold and Co., London (1931).
KRUYT, H. R., and van KLOOSTER, H. S., *Colloids*, Ed. 2, Wiley (1930).
RIDEAL, E. K., *An Introduction to Surface Chemistry*, Ed. 2, University Press, Cambridge, England (1930).
WARE, J. C., *Chemistry of the Colloidal State*, Wiley (1930).
WEISER, H. B., *Colloid Chemistry*, Wiley (1939).
    Inorganic Colloid Chemistry, Vols. II and III, Wiley (1935 and 1938).

Chapter XI

Emulsions and Foams

EMULSIONS

An emulsion is a two-phase system consisting of fine droplets of one liquid dispersed in a second liquid with which it is incompletely miscible. If the droplets be large, they will rise or fall due to density difference and will concentrate correspondingly, tending to separate as a distinct layer. If the droplets be small, they will be agitated by Brownian movement, which in turn will ultimately bring them into contact. In either case, when the particles touch surface tension tends to cause coalescence, so that emulsions are inherently less stable than corresponding suspensions of similar size and density. Hence, it is obvious that for stability there must be some factor effective in preventing contact between the particles. The most important methods of preventing such contact are: first, charge on the particle, and second, the presence around the particle of a film (cf. p. 175) which is protective and resistant, but not adhesive.

Hydrosols.

Emulsions stabilized by the presence of an electric charge on the surface of the droplets are generally called hydrosols, through analogy with the suspensoid sols and the fact that they are encountered almost exclusively with water as the dispersion medium or external phase. Agitation of immiscible pure liquids, *e.g.*, oil and water, frequently produces emulsions, but they have little stability if concentrations of more than 1 or 2% by volume are used. Whenever such a dilute emulsion consists of oil and water, it is found that the oil drops are negatively charged, as evidenced by their movement in an electric field. In Fig. 1 are shown results of Powis on the electrophoretic mobilities of the

oil globules in a very dilute emulsion, to which various electrolytes were added, the electrokinetic potentials, calculated from the mobilities (p. 194), being plotted against the millimols of added electrolyte. The valence effect, with reversal of charge when polyvalent ions were added, is clearly seen, and, furthermore, coalescence of the globules was found to occur whenever the ζ-potential fell below approximately 30 millivolts (cf. p. 200). The charge is presumably due to adsorbed ions, a fact which explains the apparent impossibility of producing emulsions of water dispersed in oil in the absence of a stabilizer. The precise source of the stabilizing ions is uncertain, but may be hydroxyl ions from the water.* The hydrosols are often quite stable, but may be broken by the methods already discussed under suspensoids, e.g., by the addition of electrolytes. The Hardy-Schulze rule holds for hydrosols as well as suspensions, di- and trivalent contra-ions being more effective than monovalent in causing loss of stability and resultant separation of the two phases.

Fig. 1. Cataphoresis of Oil Globules in Water.†

Oil hydrosols are occasionally met in industrial practice. During steam distillation, e.g., of aniline, part of the organic liquid may form an emulsion with the condensed water. Still more undesirable are the engine condenser water emulsions, which frequently contain oil emulsified in water to the extent of about 1 part in 10,000, and are extremely stable. Such a condensate is unsuitable for use as boiler feed water because the

* The KCl curve of Fig. 1 indicates preferential adsorption of the chloride anions by the oil drops at low concentrations, suggesting that the hydroxyl ions may act similarly.
† Data from Powis, F., Z. physik. Chem. 89, 186 (1914).

oil deposits on the heating surface of the boiler and interferes with heat transmission. The emulsions can be broken by precipitation with polyvalent cations, *e.g.*, alum, subjecting the hydrosol to a potential difference so that electrophoresis occurs, or by filtration through beds of such materials as calcium carbonate or active charcoal, which retain the oil, allowing only clear liquid to pass through.

Stabilized Emulsions.

Emulsions stabilized by a protective film, in contradistinction to the hydrosols stabilized by electric charges, are by far the more important from a practical point of view. The film must protect the particles from coalescence due to contact. To accomplish this, it is clearly essential that the film be in the dispersing liquid around the droplet rather than within the droplet itself. Consequently, it is not surprising to find that the film-forming protective agents are almost always materials soluble in the external liquid phase and relatively insoluble in the liquid constituting the dispersed droplets. In specific cases, *e.g.*, saponin, casein, etc., with hydrocarbon-water emulsions, the film around the drops, which prevents coalescence on collision by means of its mechanical strength, can be seen under the microscope or occasionally with the naked eye; but even if the film is completely invisible, it can prevent contact only in case it is relatively rigid and mechanically strong. Since the external liquid itself is presumably of low viscosity, relative rigidity of the film on the droplet surface can be achieved only by adsorption, in the surface layer of the liquid interface external to the drop itself, of a material capable under these conditions of setting to a gel-type structure. In general, this will require positive adsorption of the protective agent at the interface and this, in turn, usually involves an agent of high molecular weight. Finally, even assuming these conditions satisfied, if one is to avoid droplet agglomeration, the protective film must be non-adherent, so that even if two films come into contact there will be little resistance to their breaking away. This last characteristic is likely to be found only in highly solvated, emulsoidal materials. These, therefore, are the ones which are of outstand-

ing importance as the protective agents in the formation of emulsions. As typical examples for stabilization of emulsions in water may be mentioned gelatin, casein, lecithin, higher alkyl sulfonic acids, and, in particular, the soaps.

Evidence that the emulsifier does form a film around the droplets is found in the electrophoretic behavior of oil-in-water emulsions stabilized with gelatin. As the concentration of gelatin at any given pH is increased, the electrophoretic velocity at first decreases but finally reaches an approximately steady value (Fig. 2); at high pH more gelatin is required to accomplish this than at low, since the negatively charged gelatin particles are not so readily adsorbed by the oil globules as are the positively charged ones. If sufficient gelatin be present the oil globules do not move in an electric field at pH 4.7, the isoelectric point of

FIG. 2. Cataphoresis of Oil-in- Water Emulsions Stabilized by Gelatin.*

gelatin. The conclusion seems inescapable that the oil globules are completely surrounded by a film of gelatin and there appears to be no reason why this should not also be true for other stabilizers.

Preparation of Emulsions.

Emulsions are usually prepared by slowly adding the phase to be dispersed to a solution of the emulsifier or stabilizer dissolved in the other liquid, with continuous active agitation. In this way the stabilizer is readily and uniformly adsorbed at the interface between the two phases. The concentration of emulsifier is usually a fraction of 1%. In some cases excess of emulsifier actually prevents emulsification.

The conditions of emulsification, *i.e.*, the method of agitation, rate at which the second phase is added, etc., greatly influence not only rate of formation but the character of the final emulsion

* LIMBURG, H., *Rec. trav. chim.* 45, 875 (1926).

obtained.   Thus, particle size, stability, and even phase type
(W-O or O-W) * may be changed by variation of the emulsifica-
tion conditions.   Unfortunately, insufficient systematic work
has been done upon the different variables and their interrela-
tionships, making it often difficult to correlate cause and effect.

The method of agitation is important.   For any given mix-
ture continuous shaking is found to be far less efficacious than
intermittent, particularly at high vol-
ume concentrations of the dispersed
phase.   Continued agitation, intermit-
tent or otherwise, may eventually break
an emulsion, *i.e.*, there can be an op-
timum time of agitation, after which
the particle size of the emulsion in-
creases, leading finally to separation of
the phases.   For the emulsification of
benzene in 1% sodium oleate solution,
Briggs obtained the results shown in
Fig. 3, time for emulsification tending
to become infinite at large volume con-
centrations of benzene.   In practice, it
is found that emulsification time is
much reduced if the benzene is added slowly during emulsifica-
tion rather than all at once at the start.   Moderate increase in
temperature usually favors emulsification.
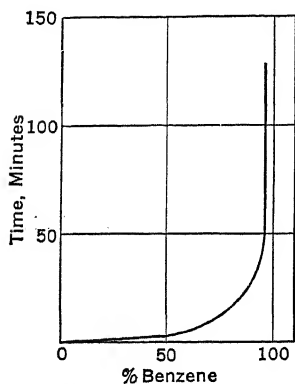
Emulsions may be made commercially with the colloid mill
(p. 113) or with high-speed stirrers; thus prepared they are
frequently coarse and of greatly varying particle size.   For these
reasons they are often submitted to a process known as homo-
genization, by which their particle size is made more uniform
and frequently reduced to a fraction of its former average value.
In the type of industrial homogenizer illustrated in Fig. 4, the
coarse emulsion is forced under great pressure through a drop-
and-lift valve, the ground surfaces A and B fitting accurately.
A strong spring C prevents A from rising more than a few

FIG. 3.   Emulsification
of Benzene in 1% Sodium
Oleate Solution.†

---

* Because water is usually one of the two liquids employed, the other frequently being
an organic liquid or "oil," it is customary to speak of oil in water (O-W) or water in oil
(W-O) emulsions, as water or oil, respectively, forms the continuous phase.

† BRIGGS, T. R., and SCHMIDT, H. F., *J. Phys. Chem.* **19**, 478 (1915).   (The Williams
and Wilkins Co.)

thousandths of an inch.   The valve opens and closes at high
frequency, which is apparently necessary for the result.   Such
homogenized emulsions are much more nearly monodisperse
and, in general, their particle size is considerably smaller than
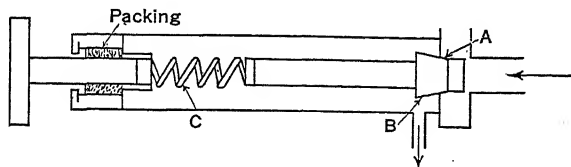
FIG. 4.   Homogenizer.

in an ordinary emulsion.   Thus, by homogenizing milk the
average diameter of the fat globules may be reduced from about
3μ to 0.2μ and the globules no longer separate as cream when
the milk is allowed to stand.   Homogenizers find extensive use
in the manufacture of margarine, which consists essentially of an
emulsion of natural fats in water or skim milk.

Phase Changes.

   With two given immiscible liquids, certain emulsifiers yield
dispersions of one liquid in the other, while other emulsifiers give
the reverse type of emulsion.   The two types may be dis-
tinguished by a variety of methods.   If the emulsion be brought
into contact with one of its component liquids the two are readily
miscible if the added liquid also forms the external phase of the
emulsion.   If, on the contrary, it forms the dispersed phase
mixing is difficult.   Alternatively, advantage may be taken of
differential solubility of a suitable third component; e.g., a dye
soluble in only one of the components may be added and, if that
component is the continuous phase, the emulsion becomes
colored, but otherwise the dyestuff simply floats on the surface
of the emulsion.   If one of the components of the emulsion is
water or other liquid of relatively high electrical conductivity,
the measurement of this property may be used as a guide to the
phase distribution, O-W emulsions naturally having a much
higher conductivity than those of the W-O type.

   Consider the changes occurring when oil is added gradually
to an O-W emulsion, complete emulsification being achieved

after each addition by shaking or agitation. At low concentrations the oil globules are far apart and do not come into contact, but as concentration increases mutual interference becomes possible. If small spheres of equal diameter be packed into a given volume as closely as possible, yet without deformation, they occupy 74.02% of this volume. It might be expected, therefore, that when the volume concentration of the oil reaches this figure inversion of the two phases would occur and a *W-O* emulsion be formed. In practice, this does not generally occur. Emulsions are seldom monodisperse, so that smaller globules of oil can pack into the interstices between the larger. Furthermore, deformation of the globules may occur. Consequently, it is not surprising to find that phase inversion or breaking of the emulsion often does not take place until volume concentrations have been reached considerably in excess of 74.02%. Indeed, Pickering was able to obtain very viscous emulsions containing 99% by volume of oil emulsified in a dilute soap solution. Nevertheless, if emulsions of a mineral oil in water with a suitable stabilizer, made with varying volume concentrations up to 75%, are allowed to stand for a considerable time, the supernatant emulsions or creams which often separate are generally found to have a volume concentration of about 74%. It appears, therefore, that the close-packing ratio is followed approximately in some cases. The relative volumes of the phases doubtless have an effect upon the type of emulsion formed, but the limits are wide and are not governed completely by the close-packing ratio.

Phase inversion may occur when a stabilizer, which normally produces one type of emulsion, is added to an emulsion of opposite type. In effect, this explains the action of a salt of a divalent or trivalent metal added to an emulsion of an oil in water, stabilized with sodium oleate or other sodium soap. Soaps of the di- and trivalent metals normally give emulsions of the *W-O* type, since with increasing valency of the metallic ion the soaps become more oil-soluble and lose their solubility in water. In this sense they are antagonistic in emulsifying tendency to the alkali soaps and other agents promoting *O-W* emulsions. Consequently, formation of di- and trimetallic

soaps in situ, by addition of suitable salts, causes the alkali soap, stabilizing the O-W emulsion, to be used up, synthesizing the antagonistic W-O type stabilizer, so that the emulsion first becomes unstable and may finally invert. Di- and trivalent metals are usually effective in causing phase inversion in the order Al > Cr > Ni > Pb > Ca.

An interesting example of phase inversion was met by Wellman and Tartar when investigating benzene-water emulsions stabilized with sodium stearate or palmitate. At low concentrations of soap (Fig. 5) normal O-W emulsions were obtained, but at higher concentrations the W-O type were formed unless the temperature was raised. Thus, the areas under and to the right hand side of each of the curves of Fig. 5 represent the stability zones of the W-O emulsions for each soap. The change from the W-O to the O-W could be repeated as often as desired by alternate heating and cooling above and below the inversion temperature. Furthermore, in accordance with the rule (p. 252) that the liquid in which the stabilizer is more soluble becomes the continuous phase, it was found that at low temperatures, where emulsions of the W-O class were formed, the soap migrated into the benzene layer, forming with it a clear solution of organogel. That the relative concen-

FIG. 5. Phase Inversion of Benzene-Water Emulsions.*

trations of the emulsifier in the two phases are important is shown by the lower inversion temperatures obtained with the palmitate. Sodium palmitate is less soluble in benzene and more soluble in water than the stearate. Therefore, the stability zone of the W-O emulsion should be correspondingly smaller with this stabilizer. Potassium laurate with only 12 carbon atoms in the hydrocarbon chain is insoluble in benzene and does not give W-O emulsions.

* WELLMAN, V. E., and TARTAR, H. V., *J. Phys. Chem.* **34**, 379 (1930). (The Williams and Wilkins Co.).

Summarizing, while the emulsion type formed depends primarily upon the character and concentration of stabilizer used, it is also influenced by the method of agitation employed, the temperature, and the volume ratio of the two phases.

Theories of Emulsification.

In seeking an explanation of the stability of emulsions one is confronted with a problem of the same type met in the case of suspensoids (Chapter VI). Brownian motion produces collisions between the particles which, because of their surface tension, tend to coalesce. Hence high viscosity (hindering Brownian motion) and lower interfacial tension between the two phases (discouraging coalescence) are both factors which encourage emulsification. Thus, two phases with an interfacial tension of under 10 dynes per cm. generally emulsify readily, while below 1 dyne per cm. emulsification is frequently spontaneous. All stabilizers are found to be materials which are adsorbed at the liquid-liquid interface, lowering its interfacial tension. Thus, Donnan was able to arrange the soaps in the order of their emulsifying power by measuring the interfacial tension between mineral oil and water containing the dissolved soap. Figure 6 shows these interfacial tensions as the homologous series is ascended. The interfacial tension is not excessively decreased until sodium laurate is reached and it is with this soap that stable emulsions are first formed. Increasing stability is noted with the salts of still higher acids. Several attempts have been made to ascertain whether a monomolecular film of the soap is formed around the drops, and in several instances, rough agreement has been obtained, the area per molecule of the soap check-

FIG. 6. Relative Interfacial Tensions of Oil and Aqueous Soap Solutions.*

* DONNAN, F. G., and POTTS, H. E., "Emulgierung," *Kolloid Z.* **7**, 208 (1910) (Theodor Steinkopff.)

ing that obtained from surface-film measurements. This led to the "oriented-wedge" theory as an explanation of the inversion of emulsions when calcium soaps were used.* The stabilizing film was pictured as in Fig. 7(a) for alkali soaps, the carboxyl groups being turned outwards towards the water while the hydrocarbon chains were immersed in the oil. In the case of soaps of di- or trivalent metals the polar head presumably possesses a cross section smaller than that of the hydrocarbon chains to which it is attached; hence, it was argued that to maintain the same orientation with a minimum of strain the oil must become the continuous phase, the drops appearing as in Fig. 7(b). This oriented-wedge hypothesis, although superficially attractive, fails to account for the change of type of some emulsions with varying concentration of stabilizer and for the temperature effects described above.

FIG. 7. Orientation at the Globule Interface.†

In addition, because the radius of an average emulsion drop is about 250 times greater than the length of the soap molecules, it appears unlikely that their mere shape could exert such a considerable effect on the relative curvature of the drops.

Solid Emulsifiers.

Certain solids, while probably incapable of materially lowering the interfacial tension between the two phases, can act as emulsifiers, forming a mechanically strong film around the droplets which discourages coalescence. The structure of a droplet of water in a water-in-benzene emulsion thus stabilized is evidently somewhat as shown in Fig. 8, the solid particles forming a protective coating for the globule. Generally speaking, the finer the solid particle size the more stable the emulsion.

It is probable that solid stabilizers present another case of hysteresis and pseudo-equilibrium. Their obvious metastability makes examination difficult and frequently leads to conflicting

* LANGMUIR, I., *Met. Chem. Eng.* 15, 468 (1916). HARKINS, W. D., BROWN, F. E., DAVIES, E. C. H., and CLARK, G. L., *J. Am. Chem. Soc.* 39, 354, 541 (1917).

† FINKLE, P., et al., *J. Am. Chem. Soc.* 45, 2780 (1923).

observations.   Thus, results may be greatly influenced by choice of the phase to which the solid emulsifier is added.   In the case of water-and-toluene emulsions stabilized with hydrophobic carbon black, better dispersion is obtained by adding the carbon black to the water with subsequent addition of the toluene than by initial addition of the stabilizer to the toluene.   This is the reverse of experience with soluble stabilizers, which are usually added to the phase in which they are the more soluble.
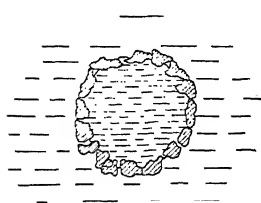
FIG. 8. Solid Emulsifier at the Interface.

The conditions under which an *O-W* or *W-O* type of emulsion are formed can be clarified by consideration of the angles of contact between the stabilizing solid and the two liquids.   Clearly, the solid must be wetted by both liquids to serve as an emulsifier, since otherwise it would not tend to stay in the boundary between the two phases.   Consider the solid particle shown in Fig. 8: for equilibrium, equate forces parallel to the solid surface (θ being the angle between the latter and the oil-water interface which envelopes the oil).

$$\sigma_{so} = \sigma_{sw} - \sigma_{ow} \cos \theta.$$

If σ_{so} be greater than σ_{sw}, $\cos \theta$ must be negative, with the result that θ is greater than 90° and the particle lies preponderantly in the oil phase. The reverse is true if σ_{sw} be greater than σ_{so}, *i.e.*, θ is less than 90°.   In order to obtain the minimum interface, droplets of the more poorly wetting liquid are formed, as shown in Fig. 8, again confirming the rule given earlier in the chapter (p. 252).

In general, both emulsions stabilized with solid bodies and those stabilized with soluble emulsifiers show analogous phenomena.   Thus, antagonism and phase inversion may also occur with solid stabilizing agents.   Carbon black emulsifies water in oil, while the reverse is true for silica.   Hence, when silica is added to the carbon black, emulsification becomes difficult or impossible.

Breaking of Emulsions.

It is often necessary to break undesirable emulsions, *e.g.*, the persistent water-in-oil emulsions found in crude petroleum and the wool fat emulsions encountered in wool scouring.   A number of methods are available.   The emulsifying agent may be de-

stroyed chemically, as by treating a soap-stabilized oil-in-water emulsion with strong mineral acid. An antagonistic stabilizer tending to induce phase inversion, if added in carefully controlled quantity, will reduce stability to the point where separation of phase is rapid. The W-O emulsions of the oil fields are often broken in this way by using O-W type emulsifiers, the salts of water-soluble sulfonic acids prepared from petroleum sludges being effective. The protective film can often be broken down by physical means; thus, freezing O-W emulsions will frequently break them, perhaps due in specific cases to dehydration and consequent embrittlement of the film by the separation of its water as ice crystals, or by the mere concentration of the emulsion as freezing progresses. Any action tending to bring the droplets into effective contact will induce coalescence and ultimate breaking of the emulsion. This can be achieved by mere agitation as in butter making, where the O-W emulsion of the cream is broken to give a mass of fat globules, the butter, and a watery liquid, buttermilk. Mere heating of the emulsion frequently induces separation and in many cases this is probably due to the same ultimate cause. Electrical methods are effective not only in O-W electrophoresis but also in W-O emulsions, the mechanism in the latter case not being entirely clear. In oil-field work a high-potential alternating current has been applied to a revolving electrode placed inside a narrow circular grounded tank. The water-in-petroleum emulsion is fed through the annular electric field continuously. Under the influence of this field the small water drops rupture the enveloping oil and coalesce. The large drops are then free to settle and separate out when the emulsion leaves the electric field.

Ageing some emulsions, particularly those stabilized with protein, causes irreversible coagulation of the protective film and results in separation of the emulsion. Soap-stabilized emulsions behave similarly on long standing, the soap separating as curds at the interface between the phases.

Detergency.

The cleansing of fabrics, utensils, and the like is a problem complicated by the fact that there are so many kinds of dirt.

Grease can be removed by solvent extraction, but this process is handicapped by the cost, losses, and toxicity and fire hazards of the solvent. From time immemorial washing has been done with water, aided by the use of agents called detergents. That their major function is the emulsification of grease and the suspension of solid particles of dirt in water, *i.e.*, action as deflocculating agents, is indicated by the fact that the oldest detergent known is fuller's earth, a highly colloidal clay of recognized deflocculating power. Because of the variations in the dirt to be removed and in the conditions of operation which must be employed, it is obvious that a general evaluation of the relative effectiveness of detergents is impossible. However, it has been found possible to standardize conditions of soiling and of washing, so that the relative results of tests parallel the practical utility of the detergent tested sufficiently to win acceptance in industry for the methods employed. Perhaps the most important method is to soil samples of fabric under controlled conditions by a mixture of oil and carbon black applied by using a volatile solvent, followed by comparison of samples after careful washing with the detergents under test, preferably using one of known effectiveness as a standard. The importance of dispersion power in detergency is attested by the fact that the capacity of a detergent to cleanse such a fabric parallels its power to disperse suspended particles, such as carbon black, manganese dioxide, iron oxide, and the like, in water. The direct determination of this dispersing power therefore offers a second important testing method.

For centuries the most important detergents have been the soaps, the alkali salts of the higher fatty acids. Their emulsifying power and the reasons for it have been discussed on p. 258. They have a similarly high dispersing power for solid particles. The importance of the molecular weight of the fatty acid and the parallelism between emulsification and detergency is attested by the fact that of the straight-chain fatty acids, lauric is the first whose sodium salt shows a high degree of detergency. If one plot either detergency or dispersing power against concentration of the soap, one finds (Fig. 9) an extraordinarily sharp maximum. Thus, whereas a 0.4% sodium oleate solution

will keep a specific manganese dioxide powder in suspension so that it passes through a given filter paper for an indefinite period, a 1.5% solution of the same soap will allow the powder to separate out as rapidly as in pure water. The cause for this maximum is not entirely clear but is probably related to the influence of concentration on the molecular structure of the soap in solution.

Detergent power as a function of concentration of the soap, including the optimum, is profoundly influenced by various factors. Thus, it is sensitive to hydrogen-ion concentration, best results with normal soaps being obtained at a pH of approximately 9 to 11.* It is also influenced by temperature, most soaps being more effective at lower temperature levels. This is probably due to the fact that soaps tend to give colloidal solutions at low temperatures and convert to true solutions at high.† In further confirmation of the part played by protective action of the soap solution in detergency, alcoholic solutions of

FIG. 9. Suspending Power versus Concentration of Detergent Employed.‡

the soap have practically no detergent action (in contradistinction to possible solvent action). Alcoholic solutions likewise show little lowering of interfacial tension against oil and no suspending power for finely divided solids such as carbon black

---

* Caustic soda alone has some detergent power (see Fig. 9). This may conceivably be due to chemical action, the caustic reacting with some of the material to form what might be termed the true detergent.

† The complex character of soap solutions is attested, for example, by the change in their appearance with change in conditions both to the naked eye and under the ultramicroscope. At low temperatures the solution is decidedly opalescent, but clears up on heating. The ultramicroscope confirms the conclusion that particle size is decreasing progressively as temperature rises. The cause of this transformation is not completely understood.

‡ VAIL, J. G., *Ind. Eng. Chem.* **28**, 294 (1936).

or manganese dioxide.   It follows, therefore, that a major factor in the effectiveness of soaps as detergents is probably the formation of protective films around droplets of grease or solid particles suspended in the water.   In other words, their action is primarily that of a protective colloid (p. 175).   Complete wetting of the solid is also always necessary before any detergent action can occur.

The complexity of the problem of detergency is confirmed by the difficulty of washing one's hands when soiled with automobile grease.   With soap alone it is almost impossible to get the hands clean, but if one will first wash the hands with a small amount of oil, preferably of low viscosity, and then with a soap solution, the grease comes off easily.   The particles of carbon and graphite are imbedded in the pores and the skin is coated with grease. The tendency of the soap solution to separate these from the skin is too low to be effective, however stable the dispersion may be once it is formed.   The solvent oil dissolves the grease and lowers its viscosity; it penetrates the pores, where it wets the carbon particles preferentially, and loosens them; hence when the soap solution is applied it can exert its full emulsifying effect. It is highly probable that soap itself, particularly in very concentrated solution, can exert considerable solvent power of this sort.   Thus, one will often obtain better results by rubbing a concentrated soap jelly into the hands before dilution than by using a dilute soap solution in the first place.

Because of the extreme complexity of the situations in which a given detergent must be used, it is not surprising to find that a mixture of detergents is frequently more effective than any one alone.   Thus, the alkali salts of many organic acids have detergent power, even though it be inferior to that of the fatty acid soaps.   One of these, rosin soap, consisting largely of the sodium salt of abietic acid, is frequently incorporated in commercial laundry soaps.   While of itself inferior to the fatty acid soaps, when mixed with them it gives more detergency for the same cost than can be secured by either alone.   Soaps of this type almost always include certain other agents, the so-called "builders."   Perhaps the most important of these is sodium silicate, itself a detergent (Fig. 9), but likewise helpful as a buffer

in control of pH and perhaps also through a mild abrasive action. Sodium carbonate and di- and trisodium phosphates are also frequently incorporated. Practically all surface waters have at least some degree of hardness, which precipitates the soaps and correspondingly destroys their detergency. One major function of the carbonates and phosphates is probably the precipitation of this hardness in a relatively inoccuous form. For special purposes, gritty materials such as ground silica, etc., are incorporated.

The higher the melting point of the fatty acid of which an alkali soap is prepared, the harder the finished cake of soap becomes and the less its solubility in water.* If a soap is too soft it dissolves rapidly and wastage in use is excessive; if too hard, it may be insufficiently soluble for satisfactory detergency. It is consequently important to control the quality of soap from this point of view. This is done by controlling the titre † of the fatty acids of which it is made. Unsaturation of the fatty acids lowers the titre and molecular weight raises it.

Soaps suffer from two serious limitations, sensitiveness to hardness and to acidity in the water employed. During recent years an important development has been the introduction of sulfonated fatty alcohols and similar compounds as detergents. The parallelism with the behavior of the fatty acids as soap-forming materials is impressive. Both have polar groups on the end of a straight hydrocarbon chain. The alkali salts of both give gel-forming solutions. In both cases the chain must be long enough to give marked depression of surface tension of the water, but not sufficiently long to make the alkali salt insoluble. Apparently the optimum chain length of the alcohols is from 10 to 14 carbon atoms per molecule. The advantage of the sulfonated alcohols is the fact that, as salts of relatively strong acids, they can exist in solutions of low pH without serious decomposition and that their calcium and magnesium salts are of much higher solubility than those of the fatty acids. They do not form suds as freely as the fatty acid soaps, but possess remarkably high detergency. Their use is spreading rapidly.

* Potash soaps are also softer and more soluble than soda soaps of the same acid.
† The initial freezing point, measured under specified conditions of cooling.

Unfortunately, adequate sources of the alcohols at low cost have not yet been developed. They may be prepared by hydrogenation under controlled conditions of the low molecular weight fatty acids of sperm, cocoanut, and palm kernel oils. In view of their value, new sources of the alcohols will certainly be developed and there is every reason to believe that these materials will become increasingly important as detergents.

In waters of exceptional hardness, *e.g.*, sea water, soap suitable for ordinary use becomes practically valueless. The difficulty can be met by using the sodium salts of the newer sulfonated alcohols or soaps of fatty acids of lower molecular weight, whose calcium and magnesium salts are more soluble and which are salted out in less degree by salts of the alkali metals. For this reason cocoanut oil soaps have long been used for such purposes.

FOAMS

Foams in many respects are analogous to emulsions, the dispersed globules consisting of gas instead of liquid. As with emulsions, in order to prepare a stable foam a third substance must be present; chemically pure liquids rarely froth or foam to a high degree. This third substance must evidently produce a surface layer differing in composition from the bulk of the liquid, *i.e.*, by negative or positive adsorption at the interface between liquid and gas. When two gas bubbles come together in a pure liquid there is nothing to prevent coalescence, but presence of a layer differing in composition from the rest of the liquid serves as a buffer, particularly if it also possesses some mechanical strength. Three per cent aqueous sodium chloride solutions yield rather unstable foams on agitation, slight negative adsorption occurring at the interface according to the Gibbs Equation (p. 67), since the surface tension is raised somewhat. More persistent foams originate, however, when aqueous solutions of colloidal materials such as the soaps, saponin, gelatin, etc., are shaken or beaten with air. In this case positive adsorption occurs, leading to a marked lowering of the surface tension at the gas-liquid interface. In general, the foam stability is greatest when the concentration is in the neighborhood of that required for maximum lowering of the surface tension: greater

concentrations make the composition of the adsorbed surface layer approximate more closely to that of the solution, with the result that the concentration difference attainable between surface layer and solution is reduced. The character of the adsorbed film at the interface has an important influence; thus, with some foaming agents, *e.g.*, saponin, rigid films are apparently formed, presumably due to the high concentration of the stabilizer induced by its adsorption in the interface. Proteins also give stiff and permanent foams and it is interesting to note that gelatin gives the most stable foam at its isoelectric point, at which point it has its lowest solubility (p. 211). Finely divided solids can stabilize foams as well as emulsions, passing into the interface and giving mechanically strong films.* There is an optimum particle size, very fine and very coarse powders being less effective than those of medium size. Solid matter becomes particularly active in foam stabilization when traces of colloidally dispersed material are also present in the liquid.

In some cases foams can be broken by the addition of a small amount of a third substance; *e.g.*, ether and the higher alcohols are often used to prevent frothing, one drop of ether completely destroying the foam on a glass of beer. Perhaps the added component is still more powerfully adsorbed at the interface than the original stabilizer, yet possesses less foaming power. The foaming of solutions during boiling may often be suppressed by blowing a current of air over the surface. This cools the vapor inside the bubbles and they tend to collapse in consequence.

Flotation.

Concentration of ores by segregation of the mineral from the gangue by hydraulic methods is extraordinarily effective when applied to mixtures of coarse particles, but rapidly loses efficiency with diminishing particle size. The disposal of metallurgical fines has always constituted a problem and for a long period the utilization of many low-grade ores was impossible. Non-ferrous metallurgy has been transformed in the last forty years by the development of ore concentration by flotation brought about by selective adhesion of gas bubbles to suspended solid particles.

---

* It is probable that finely divided solid matter often causes frothing in steam boilers.

If air is bubbled through an aqueous suspension of fine particles which are not wet by water, a concentration of particles occurs at the surface, indicated by accumulation of a slime on the vessel walls above the normal level of the liquid.   However, normally this accumulation is inconsiderable because of resuspension of the particles by the mechanical agitation of bubbling.   If, however, the liquid froths, there is an increased accumulation of the particles in the froth which under the proper conditions may become highly selective.

If the air is bubbled through an aqueous suspension of finely ground, low-grade copper sulfide ore, no significant segregation takes place, but if a suitable frothing agent be added to the water in which the solids are suspended, e.g., soap solution, a considerable percentage of the sulfide concentrates in the froth. If, in addition, a small amount of alkali xanthate be added, the concentration of sulfide in the froth can be made highly selective, leaving the gangue particles in the water, so that the process becomes an important tool in ore concentration.

Apparently in the latter case the surface of the sulfide particles undergoes incipient oxidation by the dissolved oxygen in the water; the oxidation products, chemically very reactive particularly when freshly formed, combine with the xanthate to give highly insoluble salts which adhere strongly to the surface of the sulfide.   Whether this coating of the surface is purely mechanical or is adsorptive in character is not certain.*   In any event, the alkyl radical of the xanthate is evidently outwardly directed from the surface, rendering it water repellent.   If, now, the frothing agent present in the water has a hydrocarbon radical combined with a polar group, the surface of the froth will be non-polar and possess a high affinity for the alkyl radicals with which the ore surface is coated.   Consequently, an ore particle coming into contact with a gas bubble rising through the suspension should orient itself on the air side of the air-water

---

* Thus, the xanthates may form a molecular film over the surface by chemical adsorption, e.g.,

$$PbS + S{=}C\begin{smallmatrix} \diagup OR \\ \diagdown SK \end{smallmatrix} \rightarrow PbS{=}S{=}C\begin{smallmatrix} \diagup OR \\ \diagdown SK \end{smallmatrix}$$
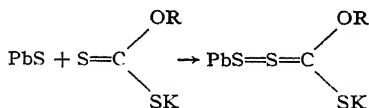
interface and, remaining there, be carried into the foam, whereas the ordinary, water-wet particles of gangue show no such tendency, and sink despite their lower gravity.

Materials which behave like the xanthates are called collectors, usually 0.05 to 0.3 lb. per ton of ore being sufficient for adequate film formation. While the xanthates were at one time the most important, many others are used, *e.g.*, thiocarbanilide, hexylamine, the dithiophosphates, and the reaction product of cresylic acid and phosphorus pentasulfide. The frothing agents most widely used are pine oil, containing terpene derivatives with hydroxyl groups, and cresylic acid, quantities of from 0.05 to 0.4 lb. per ton of ore being required. The froth after removal can be broken by jets of water and the mineral recovered.

The reactions are highly selective and can be controlled by choice of both operating conditions and addition agents employed. In this way it is possible to secure extraordinarily effective separations, not only of mineral from gangue, but selective separations of one mineral from another, as of copper sulfide from zinc sulfide, and the like. Furthermore, these separations are not limited to sulfides or heavy metal ores. In doing this, precise control of conditions is essential. Thus, maintenance of the proper pH at each point in the operation influences not only the activity of the addition agents but also the condition of the ore itself, as by control of oxidation rate. The addition of the collecting agent in carefully controlled increments can be effective in securing selectivity. Certain materials, *e.g.*, various electrolytes, lime, and the like, have distinct inhibiting action, particularly on the collectors. Others, so-called activators, definitely increase the effectiveness of the collecting agent. Thus, copper sulfate is occasionally used for this purpose. If the ore particles tend to agglomerate they are prone to entrap a considerable amount of gangue, and deflocculation of the ore, as by addition of sodium silicate solutions (cf. p. 175), can under specific conditions decidedly improve separation. Because each ore and set of conditions constitutes a specific problem which must be studied on its own merits, one encounters an infinite variety of relationships in flotation operations, but the general nature of the phenomena controlling flotation is now understood.

BIBLIOGRAPHY

BERKMAN, S., and EGLOFF, G., *Emulsions and Foams*, Reinhold (1941).
CLAYTON, W., *The Theory of Emulsions and their Technical Treatment*, Ed. 3, Blakiston (1935).
GAUDIN, A. M., *Flotation*, McGraw-Hill (1932).
INTERNATIONAL SOCIETY OF LEATHER TRADES' CHEMISTS, *Wetting and Detergency Symposium*, London (1937).
PETERSEN, W., *Schwimmaufbereitung*, Steinkopff, Dresden and Leipzig (1936).

Chapter XII

Crystalline and Amorphous States

Behavior of Solids and Liquids under Stress.

Crystals are solids of symmetrical shapes, bounded under conditions of equilibrium by plane surfaces, spontaneously assumed. The most important property differentiating a liquid from a crystalline material is that the former, unlike the latter, cannot be at equilibrium under a shearing stress.* Whereas a crystal suffers a deformation which, provided it is not excessive, is proportional to the distorting force, i.e., proportional to the square root of the work-energy of deformation (Hooke's Law), in a liquid, deformation alone does not require a finite force or expenditure of work.† Any small volumetric element of the liquid or a foreign body suspended in it can move freely throughout the mass, provided velocity of motion is low enough to make friction negligible.‡ In consequence, in a liquid at rest any increase of external pressure at a given point is transmitted throughout the liquid equally in all directions, whereas in a solid a similar pressure will develop internal distorting stresses.

Most solids under moderate stress, as, for example, under tension, undergo a deformation, which occurs practically instantaneously and is completely reversible (*i.e.*, on removal of the stress the specimen returns to its original dimensions). Such a deformation is called an elastic strain. In this sense a crystalline material, at moderate stresses, is capable of elastic strains only, whereas a liquid is subject to both elastic (*e.g.*, compression strains) and non-elastic distortions. In liquids the two types of

---

* This is true of any fluid, liquid or gas.

† Change of surface of a liquid involves the energy effect of surface tension (p. 36), and change of volume an elastic work similar to that of solids, but isothermal change of shape can be effected at constant surface and volume, in which case no energy effect is necessarily encountered.

‡ In a liquid sufficiently viscous this may be impracticable of realization.

strains are usually superimposed, though one is frequently negligible relative to the other.

Amorphous Solids.

Experimental examination of many materials shows that they exhibit the resistance to flow characteristic of the crystalline state but no evidence of a tendency to assume the geometrical form of crystals and little or no organization of structure. The molecules are evidently arranged at random and the physical properties of the solid are identical in all directions. Thus, no tendency to form plane surfaces is noticed on cleavage but conchoidal fractures may be obtained. Occasionally these materials exhibit the elastic properties of crystals; for example, over considerable ranges their extension may be proportional to the applied tension. Frequently, however, if a load, even a relatively light one, be sustained for a sufficient length of time, the material will develop a detectable permanent set, *i.e.*, will flow as though it were a liquid of extremely high viscosity. Moreover, this distortion is permanent, even though it may be small compared with the elastic or perfectly reversible deformation suffered by the same material, which disappears on removal of the load. When heated, such materials show no sharp melting point or transformation, although they do soften progressively, the tendency toward permanent set under load increasing relatively rapidly with temperature. Finally, if the temperature level be such that the material can flow under shear, it is also true that clean (preferably fresh) surfaces coalesce under sustained pressures, the magnitude of the necessary pressure and the time of application being less the higher the temperature.

Such materials are called amorphous solids. They can be looked upon as supercooled liquids. In many cases they can be prepared by cooling from the liquid state, although this is often impossible, usually due to thermal instability at the temperatures necessary to achieve a high degree of fluidity. Justification for considering them supercooled liquids is found in the fact that their flow characteristics are substantially those one would anticipate from extrapolation of the viscosity curve of Fig. 14, Chapter II. In other words, if a liquid can be cooled greatly

without crystallization, its viscosity tends to rise to a very high value, so high that flow under moderate stress can become negligible and difficult to detect experimentally.

This interpretation of the structure of amorphous solids is supported by X-ray analysis. The bending of light by a so-called diffraction grating has long been known. Thus, if a ray of visible light be passed through a glass plate on whose surface is scratched a relatively large number of parallel lines, the ray is bent, the angle of deflection depending on the distance between the lines and the wave length of the light. Study of these gratings makes it clear that the diffraction effect is dependent upon four factors: the light must pass through a region in which alternate zones differ significantly in capacity for unobstructed transmission of the light; furthermore, these zones must be approximately parallel, approximately equidistant, and the distance between zones must be of the order of magnitude of the wave length of the light in question.* Granting orderly arrangement of the atoms in a crystal, it is clear that they constitute a series of diffraction gratings, one behind the other. Presumably, the orderliness of arrangement is far greater than that of any grating scratched on glass. One would therefore expect such a crystal to function as a grating, provided one can find light of the proper wave length, far below that of visible light. X-rays of the proper range meet the requirements perfectly, and their use has made it possible to determine quantitatively the arrangement of the atoms in the structure of crystals.

X-ray analysis has also shown that crystals fall into three major classes. In the first, the homopolar or covalent crystals, the arrangement of the atoms is exactly what one would expect on the assumption that they are held together within the crystal by homopolar valence bonds. Thus, in the diamond each carbon atom is surrounded by four others, the atoms being symmetrically located at tetrahedral angles between the lines of centers. In the second class, the interatomic arrangement does not bear

* Study of the characteristics and performance of ordinary diffraction gratings scratched on glass makes it clear that the differences in light-transmitting capacity of the alternate zones need not be great, and that the zones need not be absolutely parallel nor identical in distance at all points. However, if diffraction is to be precise, the zones must be parallel and equidistant *on the average.*

a simple relation to homopolar chemical valence. Thus, in sodium chloride, around each sodium atom are found six chlorine atoms, and vice versa. One can pass through the crystal alternately from sodium atoms to chlorine, but the linkages between adjacent atoms are evidently not primary valences. The interatomic forces represent a balance of electrical attractions, the structure being characteristic of highly polar compounds which tend to ionize strongly in solution. On this account, they are called polar or heteropolar crystals. For both these types of primary valence the energy content of the bonds is computed to be of the order of 100 kg. cals., while the interatomic distance is about 1 to 1.5 Å in the homopolar compounds, 2 to 3 Å in the heteropolar. In contrast to these types of crystals, the atoms in the third class are grouped in molecules, the arrangement being that deduced from the chemical behavior of the compound. The molecules in turn are arranged in the crystal in orderly fashion, but with no evidence of normal valence bonds holding the molecules together. A characteristic example is the crystal structure of naphthalene, in which the two six-carbon rings are recognizable from the X-ray diagrams. The centers of the atoms in neighboring molecular groups approach each other much less closely in this type of crystal, the interatomic distance from group to group usually being 3 to 4 Å. Furthermore, the cohesion between neighboring molecules is small—of the order of 1 to 10 kg. cals. per group. Most organic materials fall into this third classification. One is forced to the conclusion that the molecule must be held in position in the crystal lattice by intermolecular attractions (see p. 9) or by secondary valence forces; consequently, members of this class are called secondary valence crystals.

As is to be anticipated from this analysis of their structure, primary valence crystals are characterized by high mechanical strength and extremely high melting and boiling points, whereas partial valence crystals are soft, weak, and very low in melting and boiling points. Polar crystals are intermediate in these respects. In other words, the intracrystalline forces are extraordinarily strong in homopolar and heteropolar crystals but very weak in partial valence materials.

Whereas crystals examined by X-rays yield sharp, definite diffraction patterns indicating a definite orientation of atoms or molecules, amorphous solids, on the other hand, normally yield indefinite patterns similar to those of liquids, indicating an unorganized, random arrangement, although in important cases evidence of a partial arrangement (*e.g.*, fiber pattern) is obtained (see p. 144). X-ray examination has developed into one of the most important methods of investigation of amorphous materials (see pp. 153, 162, 241, 283, 407, 495).

Many crystalline solids (particularly those existing in the form of a dense agglomerate of crystal grains, *e.g.*, the pure metals) which deform reversibly at low stresses will flow like amorphous solids at high ones. The stress beyond which permanent set can first be detected, *i.e.*, at which flow begins, is called the elastic limit. Permanent set itself is defined as the extent to which deformation fails to disappear on release of stress.

Distortion of Solids under Stress.

For normal solids under tension or compression, up to a given load, the deformation is proportional to the applied load (Hooke's Law).

$$F = E \frac{\Delta L}{L} \tag{1}$$

where L is the length of the specimen, F the load, and E the modulus of elasticity. Table I gives approximate values for

TABLE I

Substance	Modulus of Elasticity, in lbs./sq. in.
Copper	18.5×10^6
Mild steel	30×10^6
Aluminum (wire)	18×10^6
Brass	12×10^6
Glass	10×10^6
Rubber	300 [a]

[a] Varies enormously (see p. 436).

the moduli of some common solids and a number of typical stress-strain diagrams are represented in Figs. 1 and 2.

The load beyond which Hooke's Law fails to hold is termed the proportional elastic limit. At a specified load at or beyond the elastic limit, many ductile, crystalline solids, such as the metals, exhibit relatively rapid elongation without further increase in load. If this elongation continues the specimen ruptures and the maximum load is the breaking load. The elongation may slow up at constant or even increasing load and a considerable further increase may be necessary to reach the breaking load. The load at which this first elongation by flow develops and becomes observable is then called the yield point.* In addition to the initial, rapid deformation which is substantially reversible, a material loaded below the elastic limit may undergo a slow, further deformation which continues over long periods of time at a progressively decreasing rate. In many cases, when occurring well below the elastic limit, it probably ultimately ceases. The greater this slow deformation, the larger is found to be the permanent set on removal of the load. This type of deformation is called creep. The complications encountered where elastic and viscous deformation are superimposed are best illustrated by a specific case. Thus, Fig. 3 shows the variation in rate of shear of a sample of asphalt at 25° C. when subjected to a constant shearing stress of 40,700 dynes per sq. cm. Initially the rate of shear is large, but it drops

FIG. 1. Stress-Strain Curves of Some Amorphous Materials.†

---

* Note the difference between this yield point and that characteristic of fluid suspensions described on pp. 138, 219, 224, etc.

† HOUWINK, R., *Elasticity, Plasticity and Structure of Matter*, Cambridge University Press (The Macmillan Company) (1937).

rapidly to a constant value after twenty minutes.   From this
point on it remains unchanged, the asphalt behaving like a
normal liquid as long as the load is kept constant.   On removal
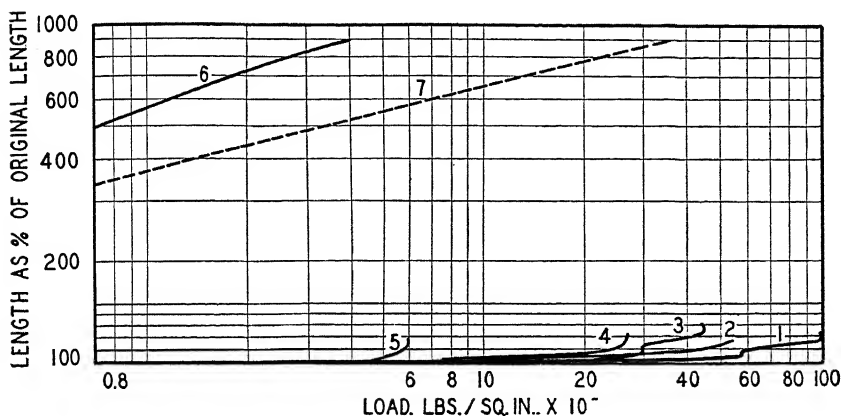
Fig. 2.   Stress-Strain Characteristics of Various Representative Ma-
terials, Plotted on Logarithmic Scales to get the Data on a Common Diagram
to Facilitate Comparison.

1. Nickel-steel.	4. Rayon.
2. Silk.	5. Wood.
3. Soft Steel.	6. Vulcanized Rubber.

7. Vulcanized Rubber, the load being expressed per unit
   area of the stretched rubber rather than unit area of
   the original cross section.

of the load, the asphalt shows a recovery approximately equal
to the excess deformation before constant rate of shear was at-
tained.   In this example the elastic effect (characteristic of a
solid) and the non-elastic one (characteristic of the liquid) can
be clearly distinguished.   The rate of distortion due to flow
is the asymptote of the first curve.   When the load is first
applied, there is an elastic distortion in excess of this flow value
(corresponding to the area ABC).   If after the asymptote is
reached the load is removed, distortion due to flow ceases, but
recovery due to elasticity occurs (corresponding to the area DEF,
approximately equal to ABC).   This case is, however, excep-
tional in that the elastic recovery is slow.

    The group of substances thus classified as amorphous solids
includes an extraordinary number of materials of great tech-

nical importance, particularly for purposes of construction. Many of them are mechanically strong, tough, extraordinarily resistant to chemical and physical action, and possess valuable elastic properties. They include leather, rubber, all the textile fibers, cellulose and its derivatives, glass, paints and varnishes, synthetic resins, and the like. On the other hand, one must not conclude that all amorphous solids have the desirable physical properties which characterize the materials just enumerated. It is undoubtedly true that the great majority of amorphous substances are useless as materials of construction.

Fig. 3. Variation in Rate of Shear with Time for Asphalt at 25° C. under a Constant Shearing Force of 40,700 Dynes per Sq. Cm.*

Manipulation of Amorphous Solids.

Certain of the naturally occurring amorphous substances can be used without materially changing their ultimate structure. Such, for example, are the textile fibers: cotton, linen, and wool. They are generally purified, colored, and mechanically fabricated into desired shapes with as little change as possible in physical structure.   However, in the great majority of cases, amorphous materials must suffer a change in plasticity before they are brought into a workable form.   Usually, they are rendered plastic by some change in condition or constitution.†   While in this plastic state the materials are molded, shaped, or worked into the desired form, and then rendered rigid and strong by destruction of the plasticity.   So important are these different methods of plasticizing amorphous materials in their industrial

* CHISWELL, E. B., JR., *Mass. Inst. Tech. Thesis* (1937).
† This initial change in plasticity may be either a decrease or an increase.

fabrication that they can helpfully serve as a basis of classification of many industries.

Certain materials, while sufficiently stiff and rigid for many uses, are none the less plastic enough even at normal temperatures to be shaped and formed.   By far the most important are those metals capable of cold stamping, spinning, and drawing, but these are characterized by changes in structure fundamentally different in kind from those underlying the behavior of ordinary amorphous plastics.   Occasionally a plastic material may be employed without change in structure after manipulation, as in the case of cold pitch for caulking cracks and seams; but usually manipulation demands a degree of plasticity intolerable in the finished article.   Furthermore, the raw material rarely possesses a plasticity proper for manipulation.   Consequently, the utilization of an amorphous material usually involves at least three of the following steps: (1) synthesis of the

TABLE II.   MANIPULATION OF PLASTIC MATERIALS

Class	Description	Examples
I	Plasticize by heat; deplasticize by cooling (Thermoplastics)	Glass, asphalts, etc.
II	Plasticize by use of solvents	Adhesives, rubber cement,
II (a)	Dissolve completely; deplasticize by evaporation	lacquers, varnishes, acetate and nitrate rayon
II (b)	Dissolve completely; deplasticize by precipitation	Viscose, acetate and cuprammonium rayon
II (c)	Swell only; deplasticize by evaporation	Paper, celluloid, smokeless powder
II (d)	Swell only; deplasticize by evaporation, followed by chemical (thermal) conversion to permanently inert condition	Ceramics
II (e)	Swell only; deplasticize by coprecipitation	Leather
III	Deplasticize by polymerization or equivalent chemical transformation	Synthetic resins, rubber, drying oils

material, (2) plasticization, (3) mechanical manipulation, (4) deplasticization.*   For purposes of classification it is desirable to focus on the technique of plasticization and deplasticization, steps (2) and (4), as done in Table II.

* Auxiliary operations are sometimes necessary, such as the de-esterification of nitrate rayon.

The simplest means of plasticization is by the application of heat. Since amorphous solids are essentially liquids greatly supercooled, workability or plasticity increases rapidly with rise in temperature. This action is reversible and an object fabricated from a heat-plasticized material, as, for example, glass, becomes rigid and strong when cooled again.*

The most important method for working amorphous solids is by the use of solvents. In general, amorphous solids swell in suitable liquids (Chapter VII, p. 143) and, whether swelling is limited or unlimited, their plasticity is greatly increased. In many cases deplasticization after working is effected by mere removal of the solvent.

Polymerization can sometimes occur progressively, building up molecules of ever increasing size. Even where the original fluidity is high, it may in this way be progressively decreased to a plasticity suitable for fabrication. After manipulation, by continuing the polymerization, usually under heat and pressure, plasticity may be destroyed and the object rendered rigid and strong. Illustrations are found in a number of the synthetic resins, of which Bakelite was an early example. In a few cases polymerization may be induced or promoted by chemical addition. The addition agent, reacting with the material, may form a bridge between two or more molecules, thus greatly increasing the molecular weight with corresponding decrease in fusibility and solubility of the substance (p. 418). The use of this means of decreasing the plasticity has hitherto been restricted largely to highly unsaturated organic substances.

* There is sometimes, however, an appreciable lag in the appearance of rigidity which must be allowed for.

Chapter XIII

Thermoplastics, Glass

A liquid has no tendency to crystallize until it has been cooled to its freezing point. The lower the temperature below its freezing point, *i.e.*, the greater its degree of supercooling, the greater its tendency to crystallize. Crystallization involves two completely independent steps: first, the formation of a crystal nucleus and, second, the growth of that nucleus once formed. Solid foreign particles in the liquid can serve as nuclei, but in a homogeneous liquid the nuclei must be synthesized from the liquid itself. Both the rate of spontaneous nucleus formation and rate of growth of nuclei once formed are functions of the degree of supercooling. Both have zero values at the freezing point, increase as supercooling increases, go through maxima, and fade away to zero values at low temperatures. The indications are that the two curves have in general the shapes shown in Fig. 1, but the values of the ordinates differ enormously from one liquid to another. Some supercooled liquids show almost no tendency to crystallize, while others can be supercooled only with great difficulty.
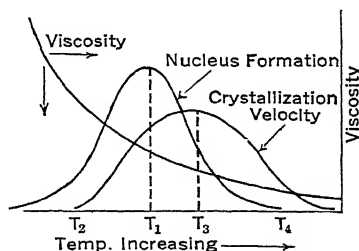
FIG. 1. Rates of Crystallization and Nucleus Formation (right hand side of diagram is the true freezing point).

The most important factor interfering with crystallization is complexity of the liquid mixture. Most pure liquids crystallize readily, whereas mixtures of a large number of substances, particularly where no one component is present in considerable amount, resist crystallization. This is obviously because crystallization can occur only when a relatively large number of

identical molecules come together in a definitely regimented arrangement, and the chance of this occurring is greatly reduced when the molecules in question are mixed with and separated by foreign molecules.* Consequently, one of the best ways to avoid crystallization of a liquid on cooling and to guarantee supersaturation, is to start with an extremely complex mixture containing no single component in considerable amount.† It seems probable that the behavior on cooling of asphalts, pitches, and tars is due to this factor. However, less complex mixtures, while they may resist crystallization, will undergo it if given time. Clearly, also, they will crystallize most rapidly if held in the temperature range from T_2 to T_4 of the accompanying Fig. 1, or still more rapidly if held at T_1 until an adequate number of nuclei are developed and then at T_3 for a sufficient period of time to allow for growth.

Another vitally important factor in crystallization is the viscosity of the liquid. Thus, the ease of supercooling of a liquid such as glycerin is probably due in part to the extraordinarily high viscosity at its freezing point, interfering with the movement of the molecules necessary for regimentation. The marked increase in viscosity of liquids as the temperature falls is probably the cause of the decrease in the rate of both nucleus formation and crystal growth on supercooling below the maxima of these curves. Unfortunately, viscosity is a characteristic of the liquid which is not easily controlled (see, however, p. 303).

Although many thermoplastics are known and find limited application (*e.g.*, sealing wax), industrially the most important are the inorganic glasses. Their use depends on the facts that they can be prepared as completely homogeneous liquids at temperatures above their freezing point, they can be held in a condition sufficiently fluid for easy plastic manipulation,‡ and

* Evidently this statement does not apply in the case of materials capable of crystallizing with each other as mixed crystals (solid solutions) or as intermolecular compounds, such as salts with water of crystallization.

† The organic chemist is always anxious to avoid this sort of result from side reactions, since it interferes with the separation and purification by crystallization of his desired reaction product.

‡ From the above discussion, it is obvious that at the temperature of plastic manipulation glasses are probably in a zone in which nucleus formation and crystal growth, while low, are none the less at or near their maximum values, so that precautions must be taken to prevent devitrification. This is all the more essential in those cases in which it is desired to reheat the once-cooled glass into the plastic range for remanipulation, because

finally, when cooled to normal temperatures where their fluidity disappears, they become rigid solids. To have widespread commercial utility a glass, when kept over a period of many years, should show no tendency to crystallization, or devitrification, as it is called. Simultaneously it should possess reasonably good mechanical properties (strength, resistance to mechanical shock, etc.) and resistance to atmospheric agencies or weathering. Furthermore, the glass should be easily workable, remaining plastic over a reasonably long temperature range—the working range—in order to permit manipulation and formation of the desired article.

Silica.

From most points of view pure silica, SiO_2, is almost perfect as a glass-forming material. In the molten condition it can be easily cooled without crystallization or devitrification. Its temperature coefficient of expansion is so small that it can be cooled rapidly without setting up serious thermal strains and the product is likewise relatively unaffected by subsequent heating and cooling (e.g., its tendency to devitrify on reheating is relatively slight). It is mechanically strong and highly resistant to chemical and physical action. It does not begin to soften until very high temperatures are reached.

X-ray studies of recent years * point to the conclusion that in fused silica at normal temperatures the whole mass consists of silicon atoms bonded together through three dimensional oxygen bridges. The divergence of the oxygen bondings around each silicon atom from tetrahedral angles is slight. The structure obviously involves three-dimensional interlocking rings, but the arrangement is not regular, the number of atoms in the rings and the interrelations of the rings varying from point to point. Such a structure makes clear the reasons for the mechanical strength and low thermal coefficient of expansion of the material. As temperature is increased, it would appear that the oxygen bridge structure becomes less stable, and the bridges progressively

during the first cooling the glass passes through a range of rapid nucleus formation and on reheating there is the risk that one will reach a zone of rapid crystal growth in which the nuclei formed, otherwise innocuous, will develop and give serious devitrification.

* WARREN, B. E., and BISCOE, J., *J. Am. Ceram. Soc.*, 21, 49, 259 (1938).

rupture with formation of oxygen-silica double bonds. Carried to the extreme this would ultimately result in monomolecular SiO_2, but the viscosity and volatility characteristics of silica indicate that any such condition is reached only at excessively high temperature levels.

Despite its advantages, silica possesses one disadvantage which has limited its use as a glass-forming material, namely, its extremely high melting and softening points, with the accompanying chemical activity in the molten state. It can be melted to a thin liquid only in an electric furnace, and in this condition it combines with practically everything sufficiently refractory to withstand the high temperature.*

Silicate Glasses.

Silica is an acidic oxide, which in the liquid state reacts with and dissolves all thermally stable basic oxides. The solutions thus produced often melt at temperatures far below that of silica itself; thus, addition of 25% by weight of sodium oxide lowers the melting point almost 1000° C. (Fig. 2). Hence, incorporation of a base offers a ready means of lowering the liquidus temperature. The compound, sodium disilicate, $Na_2O . 2 SiO_2$, gives a eutectic mixture with silica, melting at 793° C., which shows little tendency to devitrification. This glass is, unfortunately, water soluble, which makes it of little use as a material of construction. However, addition of calcium oxide in suitable amount renders a soda-silica mixture insoluble and, in certain ranges of composition, does not detract from its tendency to resist devitrification. Such mixtures are the ordinary soda-lime glasses of industry.

In glass the metallic oxides combine with the silica, causing breakdown of these oxygen bridges between silicon atoms and their replacement by polar groups with progressive change in the physical characteristics of the glass. This is incidentally the

---

* Within the last twenty years, this problem has been solved by melting the interior of a pile of sand by an electrically heated carbon resistor passing through it. The shrinkage of the mass induced by the melting of the sand next the resistor separates it from the carbon, so that carborundum formation is avoided. The viscosity of the melt prevents it from flowing down or collapsing, so that a tube of fused silica is produced, which can be further manipulated by modification of standard glass blowing methods. However, the difficulties of handling it so seriously limit its value that, for most practical purposes, the melting point of the silica must be greatly reduced.
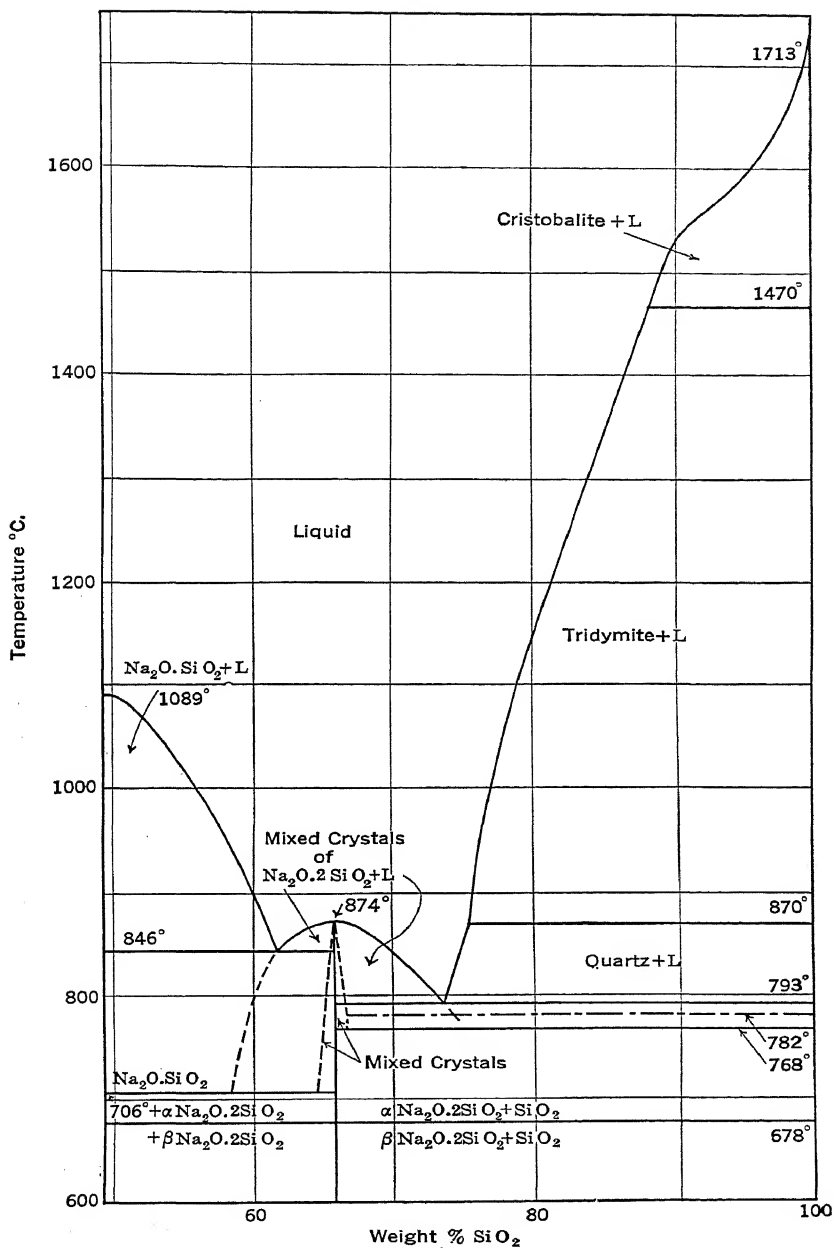
FIG. 2. Phase Equilibrium Diagram of Binary System Na_2O—SiO_2. ($Na_2O \cdot 2 SiO_2$ dissolves either SiO_2 or Na_2O as mixed crystals.)*

* MOREY, G. W., *Ind. Eng. Chem.* **25**, 742 (1933).

reason why it is necessary to limit the amount of metal in a glass if one is to conserve its valuable properties. This disintegration of the oxygen bridge structure of fused silica, breaking down as it does the three-dimensional chemical truss work of the amorphous silica itself, gives a constructive picture of the mechanism of the depression of the softening point of glasses caused by the introduction of metallic oxides.

The simple silicates, in general, do not readily supercool without devitrification and are completely unsuited for use as glass. Moreover, the range of composition of silica mixtures satisfactory for glass production is strictly limited, although composition can be greatly varied provided the variations are within these limits. To appreciate why complex silicates can produce satisfactory glasses and to understand the factors governing their utilization for this purpose, certain well-recognized analogies should be considered. Because most oxy-acids form complex metallic salts, the tendency being least in the case of the strong acids but increasing progressively as the acids become weaker (p. 184), it is not surprising to find that the weak silicic acid, containing four replaceable hydrogen atoms, has a marked tendency to form polyacidic derivatives, e.g., the di- and trisilicates. Hence, use of mixes rich in silica and containing several bases would be expected to increase molecular complexity and hence to encourage resistance to crystallization.

As already pointed out, the value of a glass depends greatly on its properties in the temperature range below its freezing point. The influence of composition on the latter is therefore of primary importance. Because the mixtures in question are complex, graphical representation is difficult, but in the case of ternary mixtures the effects are readily presented on the familiar triangular diagrams.

Because in any three-component mixture the per cent of the third component is necessarily determined by those of the other two, the composition of any such mixture can be represented by a suitable point within an equilateral triangle, in which the vertices correspond to the pure components (in the case of Fig. 3 to Na_2O, CaO, and SiO_2). The sides of the triangle are graduated in percentage of the respective component and any points on a straight line drawn through a vertex represent mixtures in which the ratio of the components shown at the other two vertices remains constant. The

location of the composition of a given mixture is determined by the intersection of straight lines through the percentage of each component, parallel to the side opposite the vertex representing that component. Thus, the point A on Fig. 3 represents 20% CaO, 15% Na_2O, and 65% SiO_2.

Figure 4 shows a part of the three-dimensional figure representing the system Na_2O—CaO—SiO_2, temperatures being represented on the vertical axis, while the composition of a mix is represented on the base. In this figure the locus of points representing melts in equilibrium with their respective solid phases forms numerous curved surfaces which intersect to form a number of valleys and sinks. The valleys represent monovariant conditions of the system or binary eutectics, and the sinks represent non-variant conditions or ternary eutectics. Figure 5 shows the projection of these valleys onto the base of the

100% Ca O

50%   50%

100% Na_2O 50%   100% Si O_2
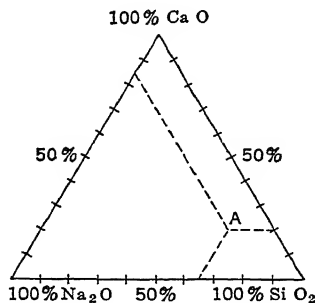
FIG. 3.   Composition on a Ternary Diagram.
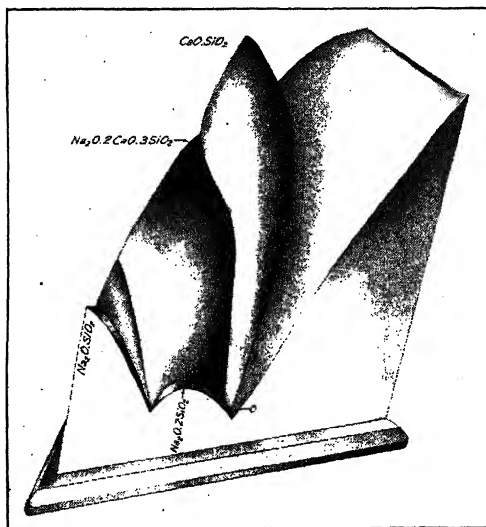
FIG. 4.   Sketch of Solid Temperature-Composition Model Representing Phase Equilibrium Relationships in Portion of Ternary System Na_2O—CaO—SiO_2.*

figure.   The larger letters indicate the composition of the various solid phases existing under equilibrium conditions in the several fields; the smaller letters indicate points corresponding to the composition of specific crystalline compounds.*   Figure 6 shows, in addition to the eutectic paths, the projections of curves defined by the intersections of the indicated isothermal surfaces with the equilibrium surfaces.   As a general rule, the composition of a
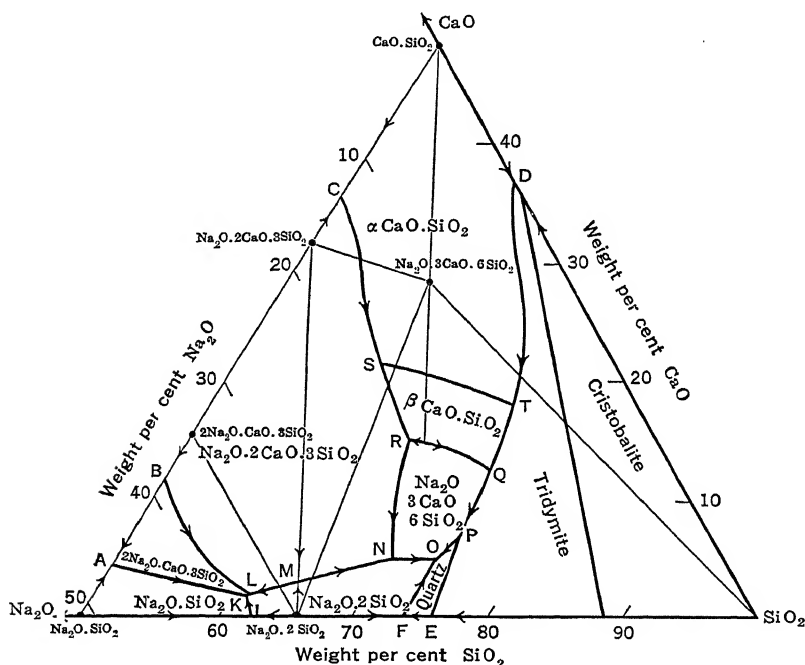
FIG. 5.   Phase Equilibrium Diagram of Portion of Ternary System Na₂O—CaO—SiO₂, Showing Boundary Curves and Tie Lines.†

melt in equilibrium with its solid phase will upon cooling follow a straight line connecting the original liquid composition with that of the solid phase, until the line intersects a binary eutectic line.   From this point of intersection the liquid composition follows the eutectic line in the direction of decrease of tempera-ture (shown on Fig. 5 by arrowheads), yielding simultaneously

---

* These may lie outside the field of liquid composition with which they are in equi-librium.
† MOREY, G. W., *Ind. Eng. Chem.* **25**, 742 (1933).

the materials corresponding to the adjoining areas, until complete solidification occurs upon reaching a ternary eutectic. Analysis of the diagrams brings out certain important points. In the first place, it is possible by the use of lime and soda to reduce the melting point of silica from 1700° to below 750°, *i.e.*, over 900° C., but in order to melt this low, the liquid must contain over 20% Na_2O and less than 5% CaO. Such a liquid will
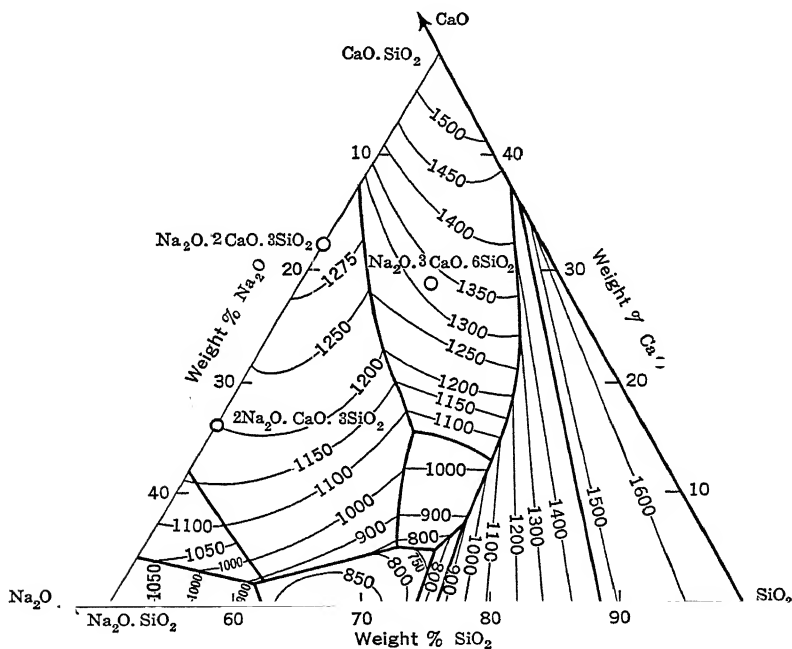
FIG. 6.   Phase Equilibrium Diagram of Portion of Ternary System $Na_2O—CaO—SiO_2$, Showing Isotherms.*

be very viscous at its freezing point, because, while the introduction of the bases into the molten silica greatly reduces its viscosity at constant temperature, the viscosity of the resultant liquid is still further increased to a very high value before its freezing point is reached, *i.e.*, before even a tendency to crystallization can develop. Furthermore, the lime and soda exist in the melt combined with the silica probably in a large number of different ways, so that tendency to crystallize is reduced by the

* MOREY, G. W., *Ind. Eng. Chem.* 25, 742 (1933).

complexity of the liquid mixture as well as by its viscosity. Consequently, the liquid is capable of cooling to a glass. In practice, it is found that the compound $Na_2O.3\ CaO.6\ SiO_2$ crystallizes only with the greatest difficulty. However, a low-melting glass of composition to deposit this material would have the disadvantage of being rich in soda and poor in lime, approaching the composition and properties of sodium silicate, water glass. While not water-soluble as such, it would be readily attacked by water and aqueous solutions generally. Water resistance can be greatly increased by reduction of the percentage of soda and increase of that of lime, because the latter forms with silica no water-soluble compounds. However, inspection of the diagram makes it obvious that any such change results in marked rise in freezing point. Experience shows that the soda can be reduced to about 15% if the lime be increased to a little over 10%, and the freezing point will still be about 1000° C. Most industrial soda-lime glasses have compositions lying near the line PQ of Fig. 5.

Introduction of potash instead of soda similarly lowers the melting point but the system K_2O—CaO—SiO_2 is much more complex, more ternary compounds being formed. Potash glasses are more resistant to weathering and devitrification and give more viscous melts.

Just as the introduction of either lime or soda to a mixture of silica with the other lowers its freezing point, so the initial introduction of additional components into the melt frequently still further reduces the freezing point, even though the addition of larger quantities may ultimately raise it. Such additions obviously tend to retard crystallization of the supercooled liquid, both because of the increase in complexity of the melt and the increase in viscosity, resulting from the lowering of the freezing point. The materials added must be soluble in the liquid and it is very desirable to keep the silica as high as possible, because it greatly increases the resistivity of the glass to the disintegrating action of aqueous solutions. Furthermore, as is apparent from Fig. 6, at high percentages of silica the crystalline material which tends to separate out at the freezing point of the melt is a polymorph of SiO_2, and there is apparently quite

definite resistance of silica to crystallization, particularly at low temperatures. This fact constitutes a second advantage in keeping the silica high. Consequently, the ratio of total bases to silica should be kept relatively constant, *i.e.*, incorporation of additional bases should be a replacement and not an addition.

In practice, the glassmaker never deals with a pure soda-lime-silica system, since, even if he wished to do so, the refractory containers are always attacked to some extent by the melt. Thus, alumina and magnesia are difficult to avoid, while iron oxide is also a common impurity, objectionable in most glasses because of discoloration.

Glass Melting.

Glass manufacture consists of two steps, preparation of the molten glass from the raw materials and fabrication of the melt into the desired articles.

Two methods of manufacture of molten glass are employed, usually depending on the scale of production. For small amounts, the glass is always melted in pots, which are in effect large crucibles, a number of which are generally heated in a single furnace. These pots may be either open or closed (see Fig. 7); the latter, with the mouth of the pot projecting out of the furnace, are used for glasses injured by contact with the flame, such
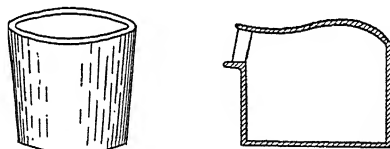
FIG. 7. Glass Pots.

as lead glasses, in which the lead is reduced by gas deficient in oxygen or converted to sulfide by the sulfur in the fuel. The constituents are introduced into the hot pot, the temperature is raised and held at about 1400° C. for closed pots and somewhat hotter for open, for melting to a completely homogeneous mass. Silica is introduced as quartz sand or quartzites; the alkalies, as carbonates or sulfates; lime, in one of its pure natural forms, such as chalk or limestone, which may be dolomitic where magnesia is desired in the glass; lead, as litharge (PbO) or red lead (Pb_3O_4). The carbonates evolve CO_2, that of lime decomposing thermally before reacting with silica. Sodium sulfate must be reduced by carbon in some form, such as charcoal or powdered

coal. Most of the reactions evolve gases, SO_2, CO_2, and CO, which are effective in stirring the mix. Red lead decomposes with evolution of oxygen, effective not only for agitation but also as an oxidizing agent. Broken glass waste (cullet) from previous batches is also introduced and serves to dissolve and dilute the alkalies, which melt first, thus protecting the pot from their fluxing action. For complete homogeneity one can granulate the melt (as by quenching in water), mix thoroughly, and remelt.

After the mass is thoroughly molten, it is essential to hold it a considerable time at a temperature level well above its freezing point, an operation called "refining." A scum of insoluble materials, "glass-gall," such as undigested particles of silica, sloughed-off clay from the refractory, etc., rises to the top and can be removed mechanically. The gas bubbles rising and escaping from the surface help in dragging up the scum. Unfortunately, the finest gas bubbles rise extremely slowly through the viscous mass and some form of agitation is necessary to eliminate them and thus prevent "seedy" glass. Most effective is provision of a stream of large gas bubbles which, on rising, will agitate the mass and sweep the fine gas bubbles out with themselves. Formerly, in this "blocking" process, the pots were stirred with green wood, the vapors from which served the purpose. At present ammonium nitrate is often used, lumps of the salt wrapped in wet paper being thrust into the pot. It is even better to evolve the gas from within the melt itself because this supersaturates the glass with gas, which, escaping into the finer bubbles, enlarges them and facilitates their elimination. This can be achieved in some degree in soda-lime glasses by the use of sulfate as a source of soda. This material is reduced to SO_2 and CO by carbon at a relatively late stage in the melting operation, thus giving the gas evolution when most needed. The reduction is prolonged over a period of time by the fact that the solubility of the sulfate in the glass is only a few per cent, the excess floating on the melt as so-called "salt water" and dissolving relatively slowly. Arsenic may be added as arsenious oxide, which probably dissolves as arsenate or arsenite. Some decomposition occurs with evolution of oxygen, which, since arsenious oxide has the advantage of being heavier than the

glass, rises through the whole mass.   The elimination of bubbles is sometimes called "plaining."

To reduce color the iron, present as an impurity, is oxidized from the green ferrous condition to the less intense, light-yellow ferric state.   In good glasses the yellow must be neutralized by addition of the complementary amount of pink color.   Manganese dioxide has long been used, but has the disadvantage of turning purple on long exposure to sunlight.   Selenium is now preferred.

Scum or gall can be removed from the pot mechanically or can be held back by a "floater" of clay or by forcing the glass to flow under a partition behind the mouth. The temperature of refining is high, to reduce viscosity; for working and blowing the

FIG. 8.   Glass Furnace.

glass, the temperature is lowered to a suitable point, depending on the composition.

For mass production of glasses which can tolerate direct contact with the flame, the so-called tank furnace, operating continuously, is employed (see Fig. 8).   The charge is fed in at the rear of the furnace, on top of the molten glass, exposed to maximum heat to get rapid melting and the gas agitation resulting therefrom.   The molten mass then flows slowly down the length of the furnace and is held at a high temperature, the flow being regulated to give ample time for separation of gas bubbles and gall.   The zone in which this refining operation occurs is often separated from the melting zone by a refractory wall, under which the glass flows through a throat or orifice.   The glass enters the lower temperature working zone under a series of clay floaters to hold back the scum; from it the glass flows to the machine or is removed through working ports.   Tanks are about 18 ft. wide, the glass depth being 3 to 5 ft., while, to give the time necessary for melting and refining, the length of the
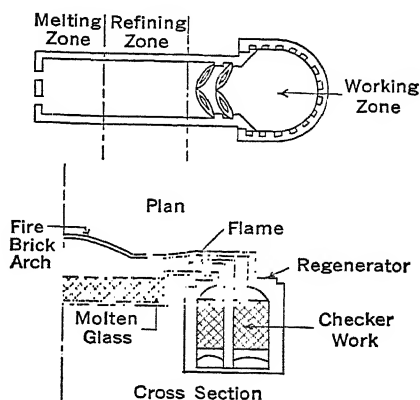
furnace is proportional to the rate of production. For plate glass, where, to attain homogeneity, plaining is of the utmost importance, very long tanks are employed. The usual temperature of operation is about 1500° C., for, although high temperatures obviously increase furnace capacity, they also seriously increase disintegration of the refractories, thus cutting down length of run and contaminating the glass.

The furnaces are fired with oil, or with producer or natural gas. With producer gas, both air and gas are preheated, preferably by regenerators located below the furnace. With natural gas, to avoid decomposition of the methane air only is preheated. The furnaces are inefficient, only 15 to 20% of the heat generated by the fuel passing usefully to the glass, the remainder being lost, mainly by radiation. In spite of this the cost of fuel is a minor factor in glass production.

Refractories for Glass Manufacture.

High temperatures are desirable in the glass-melting operation since viscosity is reduced, facilitating reaction and elimination of gas. However, the container must possess both relatively high mechanical strength and adequate resistance to the fluxing action of the molten glass. In consequence, the refractory has always set the limit to the practicable temperature of the glass furnace. While marked improvements, due primarily to purer and better raw materials together with improved methods of manufacture, have been made throughout the last half century, only in the last few years have outstanding advances resulting in progressively increasing furnace temperatures been made.

In the past, clay was the material exclusively employed for refractories (see p. 466). The diagram of Fig. 4, p. 457, for the binary system, Al_2O_3—SiO_2, shows that above 1545° C. clay begins to melt, separating solid mullite, $3 Al_2O_3 \cdot 2 SiO_2$, and a liquid phase which increases in amount with both SiO_2 content and temperature. Hence it is important to choose a pure clay with the highest possible ratio of alumina to silica. Such clay is made into bricks for the construction of furnaces, and molded wet into pots for the batch-melting of glass mixes. To secure maximum density and mechanical strength at the high tempera-

ture of the furnace, the refractories are fabricated, dried and burnt with greatest care.   Particularly in the case of large glass pots, thermal strains are so serious that the temperature must be raised and lowered very slowly and the pot, once introduced into the furnace, never allowed to cool.   At best, breakage is considerable, and the refractory always sloughs off into the glass to a certain extent.   If this sloughing amounts to more than the alumina tolerance of the glass, or if time be not available for homogeneous distribution through the melt, "stones" of undigested clay particles are found in the finished product.

The high melting point of Al_2O_3 and its low solubility in silica make it attractive as a glass refractory.   However, at high temperatures alumina and silica form the high-melting, stable compound, mullite, $3\ Al_2O_3 . 2\ SiO_2$, mechanically strong and resistant to thermal shock and fluxing action.   It can be made from kaolin enriched with alumina, by melting the raw materials in the electric furnace and pouring the product into molds, or crushing the molten mass, with subsequent molding and sintering, thus securing a density hitherto unobtainable.   While this important forward step must not be underestimated, the expense of the product has enabled it to replace the older clay refractories only in those portions of the furnace subjected to severe punishment.

Fabrication of Glass Articles.

For many centuries articles of diverse shapes and intricate designs have been made from glass by the remarkable art of blowing.   In almost every case the first step was the production of a hollow globe or cylinder of glass by air pressure from the lungs, large articles requiring great physical strength and endurance.   Within the last twenty-five years, however, ingenious machines have been developed that have largely replaced the highly skilled and clever workman of the past.   A most important property of glass in its fabrication is its viscosity-temperature relationship, since the glass must remain plastic for a period sufficiently long to allow working.   Glass is refined at a viscosity of about 100 poises, can be gathered at about 1000, blown at 1,000,000, and annealed at 10^{12} to 10^{14} poises.   For hand

blowing a long working range is not so important as for machine-made glass, since the manual worker can, to some extent, adjust the process to fit the characteristics of the glass.   For machine-made glasses a suitable working range must be obtained by altering the composition of the mix.

The general method of hand "working" is illustrated in the manufacture of window glass.   The glass, properly refined, is cooled to the range of working plasticity, and a "gob" caught on a "blowpipe," a piece of iron tubing some 5 ft. long.   By repeated insertion into the mass of glass, with intervening cooling, as much as 15 or 18 lbs. of glass may be gathered on the end of the pipe.   By blowing through the pipe and continuously rotating the whole, a hollow globe of glass is formed.   This is reheated in a furnace called the "glory hole" and by skillful coordination of blowing, turning and swinging (the workman standing on the bridge over a pit) a cylinder from 3 to 4 feet long and 12 or more ins. in diameter is formed.   When the walls of the cylinder have attained the proper thickness, it is cooled until rigid and detached from the pipe by touching the hot glass with a piece of wet wood.   The two ends are cut off, the resultant open cylinder cut lengthwise with a diamond and placed on a flat surface in a furnace where, when sufficiently soft, it opens out and flattens to a plane sheet.

In the Lubbers machine, the earliest developed for the purpose, a cylinder thirty inches or so in diameter is drawn upward, by a rising mechanical head, through an annular orifice in a refractory plate resting on the plastic glass surface.   Air pressure is supplied and the glass is cooled at a rate to form a perfect cylinder, drawn to a height of 30 or 40 ft., which is then manipulated as in hand blowing.   The newer Fourcault machine withdraws a flat ribbon of glass continuously from the plastic mass, speed and rate of cooling controlling the thickness.   The sheet is started by a pair of feed rolls dipping under the surface of the molten glass, which lift a portion and allow enough immediate cooling to enable the sheet to sustain its own weight.

Plate glass is made by casting molten glass upon a flat cast-iron plate or table, getting uniform thickness by a heavy, hot cast-iron roller traveling over the sheet, kept in position by

guides at the edges of the table. The carefully annealed plate is finished by grinding to plane surfaces with rotating discs, fed first with sharp sand, later with emery powder, and finally with putty powder or rouge.* The surface is softer and more subject to abrasion than the "fire surface" of ordinary glass.

Bottles, fruit jars, and hollow ware in general are made by blowing a small globe, placing it in a two-part mold, and again blowing to force the glass to fill the mold. The end attached to the blowpipe is cut off, the article reheated and finished. The process was formerly manual, but machines are now available that accept the molten glass at one end and turn out the finished article at the other. Dependent on the shape of the article, either air pressure or plungers may be used in the mold. Electric lamp bulbs are machine made by a pure blowing process (no molding), gobs of glass being fed automatically to air jets, which then blow them to the desired shape. Such machinery is entirely automatic, 300 to 500 lamp bulbs being produced per minute.

"Cut glass" is made from a heavy potassium-lead borosilicate glass, of high index of refraction and wide dispersion. The article is cast with deep grooves and facets. It is ground and given a high polish with putty powder fed to wooden wheels, the edges of which are shaped to the surfaces.

Annealing.

Visualize a freshly formed sheet of glass as it starts to cool. While its viscosity is high relative to that of ordinary liquids, it is still low enough so that internal strains are quickly relieved.† Because it is losing heat, there is a temperature gradient through the sheet, *i.e.*, it is hot on the inside and cooler on the surface. As it continues to cool, it reaches a temperature level at which the viscosity becomes sufficiently great so that further rate of relief of internal strain becomes negligibly slow. Granting that the temperature gradient through the sheet has not decreased significantly, no internal thermal strains have as yet been set up, but when the sheet reaches room temperature and stops

---

* The thickness may be reduced by such treatment from 30 to 50%.
† The internal flow necessary to relieve thermal strains is very small.

cooling the temperature gradient through it disappears. Since the time at which the sheet cooled below the temperature level of rapid internal adjustment, the interior of the sheet has dropped in temperature more than the exterior. Consequently, the interior of the sheet has contracted more than the exterior, thus submitting the interior to tension and the exterior to compression. In other words, the glass is now subject to an internal mechanical strain. One can follow the strain in the glass and its relief by measurements of birefringence. (See also p. 406 *et seq.*)

From certain points of view, this condition of strain represents an advantage. Suppose, for example, the sheet be bent. This increases the compression on the concave surface, but decreases it on the convex. Further bending will ultimately put the convex surface under tension. However, glass is much stronger in compression than in tension; consequently, glass in this condition will stand more flexing than were the strain relieved (so called quenched or toughened glass). Unfortunately, glass in this condition is unstable in the sense that once rupture is initiated, *e.g.*, by a scratch, the whole piece flies into fragments (Prince Rupert drops). In order to get a stable glass the internal strains must be relieved by the process of annealing.

Glass is annealed by holding it at a substantially constant temperature level selected to give a viscosity such that, while rate of relief of strain by internal flow is very slow, it is still appreciable. Constancy of temperature of the sheet eliminates temperature gradient through it. The time is sufficient to eliminate strains under these conditions. The sheet is now chilled quickly into a region where, because of increased viscosity, further adjustment of strain no longer occurs. The temperature gradient and the corresponding internal strain present during cooling both disappear when the sheet is finally cooled.*

Annealing is carried out by passing the glass through a "lehr" or annealing oven, usually continuous, the stock traveling on an endless table.

---

* The correctness of this interpretation is confirmed, for example, by the fact that the effectiveness of annealing is unimpaired by rate of cooling, provided this is not rapid enough to cause rupture by the internal strains of excessive temperature gradient through the sheet.

Since annealing is an adjustment due to actual internal flow of the glass, one would expect the rate of relief of strain to be proportional to the strain and inversely proportional to the viscosity of the glass. The data indicate the correctness of this assumption.* Analysis of the mechanism of annealing in the light of the viscosity-temperature characteristics of liquids (p. 29) makes clear the reasons for the fact that for a given glass the temperature range essential for successful annealing is quite narrow. For methods of measurement of the extremely high viscosities involved see pp. 23–24.

There are certainly other factors involved in annealing. Thus, the density of glass goes up very significantly on annealing, increasing by about 0.005 to 0.01. It seems plausible that the temperature range in which abnormally rapid increase in viscosity occurs is also the temperature range of maximum rate of conversion of double bonded oxygen linkages into oxygen bridges between silicon atoms (see p. 283). Rapid cooling from high temperature would reduce the opportunity for these changes and would be expected to result in a somewhat lower density of the glass. Since the bridge structure is evidently the more stable form at low temperature levels, preventing its formation would result in a metastable condition of localized internal strain (different in kind from thermal strains described above) which may perhaps be the cause of the fragmentation of unannealed glass on rupture.

Colored and Other Special Glasses.

Glass may be colored by dissolved metallic bases or by suspended solid particles of colloidal dimensions. That electrolytes dissolved in glass ionize is indicated by the colors they impart. Thus, ferrous and chromic oxides give green and cobalt blue.† The effect of suspended solid particles is determined primarily by particle size. Thus, the addition of a trace of gold chloride to the molten glass leaves it colorless or yellow on cooling, but on reheating it develops the deep bluish red of gold ruby

---

* Lillie, H. R., *J. Am. Ceram. Soc.*, 19, 45 (1936).
  Littleton, J. T., *Ibid.*, 17, 43 (1934).
† This blue is so intense that if the percentage of cobalt be high the glass may be ground as the pigment smalt.

glass. Overheating changes it to a dull brown by reflected light, and blue by transmitted light. The color is evidently due to colloidal gold in the glass (p. 117). As first formed, the particle size is so small, owing to the great dilution of the gold, that the color effect is negligible. On reheating, coagulation or agglomeration of the particles occurs, developing the colloidal color effect. Overheating increases the particle size and correspondingly lessens the color intensity, particularly of the blue and red components. Copper ruby is a similar effect obtained by the use of cuprous oxide, Cu_2O, which apparently dissolves at high temperature but is insoluble at low, or perhaps is actually reduced to the metal. Here again reheating is necessary for color development. Selenium oxide develops red without the necessity for reheating. Opaque, colorless glasses are made by the use of fluorspar, cryolite, or calcium phosphate as bone ash. Excess of oxides of tin, zinc, or aluminum produces a similar effect though in less degree. Whereas formerly the "opal" glasses were made by the use of melts in which the insoluble materials separated spontaneously on the cooling of the glass,
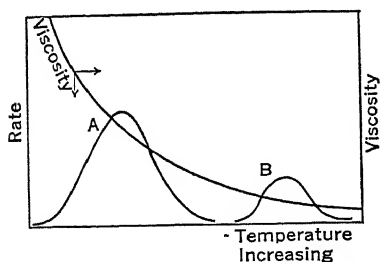
FIG. 9. Rates of Crystallization and Nucleus Formation.

it has been found possible to develop melts in which the rates of crystal growth of the opalescing constituents are given by a curve such as B in Fig. 9,* while rates of nucleus formation are represented by curve A of the same figure. By cooling the glass into the range under A for a controlled length of time, one can induce the formation of a definite number of nuclei; by now reheating into the range B, each of these nuclei grows to a crystallite of substantially uniform size, in contradistinction to the obviously non-uniform particle size of the crystallites obtained when the crystallization and nucleus-formation curves are partially superimposed, as in Fig. 1, so that formation of new nuclei and crystal growth of both old and new go on simultaneously, necessarily resulting in diversity of size of the final particles.

* Compare this with Fig. 1.

Uniformity of particle size makes it possible to secure maximum light dispersion with minimum absorption, as well as uniformity of dispersion effect. The effectiveness of reheated gold and copper ruby glasses is probably to be explained in a similar way. More recently pastel shades for tableware have been produced by some of the rare-earth oxides.

The surface of finished articles can be roughened and rendered translucent by sandblasting, etching with hydrofluoric acid, or chipping. Patterns can be secured in the blasting operation by the use of stencils or in etching by coating with wax, through which the design is cut. To avoid weakening the glass, etching is best done with dilute acid in stages. Chipping consists in coating the glass with a glue solution, the shrinkage of which on drying causes small chips of glass to be torn away from the surface, making it translucent.

Shatterproof or laminated glass usually consists of two layers of glass between which is sandwiched a layer of a plastic, such as polymerized vinyl ester. Nitrocellulose or cellulose acetate was formerly used.

Historical Development.

Since soda is the cheapest and most abundant alkali and lime the corresponding divalent base, it is not surprising to find that soda-lime glasses are industrially the most important today as they were 4000 years ago. At that time it was impossible to realize high temperatures under controlled conditions, and raw materials were relatively impure. Hence, samples of early glass analyzed (see Samples 1 and 2, Table I) are high in alkali, low in lime, and contain a relatively large number of adventitious impurities. In time the glassmakers learned how to maintain higher operating temperatures and were able to reduce the alkalies, substituting some lime by magnesia (see Samples 3 and 4 in Table I). Early in the nineteenth century purer raw materials became available, and the composition of the glass was simplified accordingly. Thus, Samples 5 to 9 inclusive are ninteenth century, hand-blown, cylinder window glass of American, Belgian, English, French, and German make, respectively. With the introduction of machine-made glasses, changes

TABLE I. COMPOSITION OF VARIOUS GLASSES *

Constituents	1	2	3	4	5	6	7	8	9	10	11	12	13
SiO_2	65.95	66.26	68.48	67.44	72.26	69.48	71.4	69.65	72.68	71.74	71.82	72.4	80.75
Al_2O_3	2.49	3.26	0.70	2.98	{1.42	{2.59	{1.90	{1.82	{1.06	1.84	1.40	0.8	{2.20
Fe_2O_3	0.28	0.78	0.91	0.51						0.14	0.07	0.4	
CaO	6.89	7.09	5.71	4.80	13.34	13.40	12.40	13.31	12.76	14.16	8.9	5.3	0.3
MgO	1.37	1.48	5.28	5.64	0	0.26	—	—	0.26	—	—	3.7	
Na_2O	20.3	19.33	14.95	13.94	14.01	14.55	15.0	15.22	13.25	11.96	3.31	17.4	4.10
K_2O	0.96	0.37	2.83	1.93	—	—	—	—	—	—	14.27	—	0.10
MnO	0.97	0.61	—	0.70	—	—	—	—	—	—	—	—	(12.0 B_2O_3)
CuO	—	0.95	—	—	—	—	—	—	—	—	—	—	(0.40 As_2O_5)
SO_3	1.08	—	0.54	0.84	—	—	—	—	—	0.29	0.31	—	—
PbO	—	—	0.95	1.01	—	—	—	—	—	—	—	—	—

* SHARP, D. E., *Ind. Eng. Chem.* **25**, 755 (1933).
1. Colorless glass from Elephantiné, Egypt 200–100 B.C.
2. Dark blue glass ditto.
3. Window glass about A.D. 900.
4. White blown glass about A.D. 900.
5–9. Hand blown cylinder window glass—nineteenth century.
10. European machine blown cylinder glass—1930 (Lubbers).
11. U. S. flat drawn window glass—1929 (Fourcault).
12. U. S. machine glass (electric light bulb).
13. Pyrex 702 EJ.

in composition designed to give greater freedom from devitrification and broader working-temperature ranges became necessary. With the semi-automatic, machine-blown window glass (Lubbers), since the process was essentially an imitation of hand blowing, no considerable changes were necessary as shown by Sample 10 of Table I, although, owing to the higher furnace temperature attainable, a slight further reduction in soda and increase in lime were possible. Attention is called to the remarkable similarities in composition of all glasses up to this point, despite diversity of both sources of raw material and conditions of manufacture.

When modern processes of drawing window glass directly as a flat ribbon were developed, it was found desirable to work at somewhat lower temperatures to secure lower rates of cooling and hence less rapid increase of viscosity with time; this was achieved by raising the soda and lowering the calcium, together with additions of small amounts of either alumina or magnesia, constituents which apparently flatten the viscosity-temperature curve (Sample 11). With the development of automatic machinery for blowing and pressing into molds to make bottles, tableware, and similar products by mass production methods, it became necessary to carry this modification of the composition still further, resulting in a glass such as Sample 12. This adjustment of composition to meet mechanical requirements of manufacture is still under way.

While the soda-lime glasses just discussed are technically important, constituting by far the largest volume of glass manufacture, it is possible to produce glasses differing widely in composition and in a sense in type, although the fundamental behavior remains unchanged. Thus, using lead instead of lime, and potash for soda, the lead-potash glasses, characterized by high refractive index and high dispersion brilliancy and used for optical purposes and the manufacture of cut glass, are obtained. Other alkalies are not frequently found in glass because of expense, but as divalent metals one can use ZnO, BaO, FeO, CdO, etc. The use of MgO has already been mentioned. Complex mixtures containing relatively small amounts of a number of these constituents are frequently found, the purpose probably

being to lessen tendency for devitrification, partly by lowering the freezing point and partly by increasing the complexity of the liquid mixture.

Boron trioxide is often added in small quantities (1%) to window glasses since it improves plaining rate, strength, weathering resistance, and brilliancy. However, a more fundamental change in the character of the glass is produced by using a relatively large amount of boron trioxide (see Sample 13, Table I), simultaneously lowering the proportion of bases. Because of this reduction, the melting point is raised, but the characteristic properties of silica as a glass are exhibited to a larger extent. Thus, the boro-silicate glasses show very low thermal coefficients of expansion, in contradistinction to the ordinary glasses containing silica as the only acidic element. They are correspondingly resistant to thermal shock, *i.e.*, sudden change in temperature. Where the amount of the basic oxide is adequately reduced, they are also more resistant to the action of aqueous solutions. Sample 13 is an analysis of Pyrex, which is representative of this type. Water resistivity can still be secured with higher alkali content and consequently lower melting point, by the introduction of divalent metal. For this purpose ZnO was formerly employed in Jena boro-silicate laboratory glassware. However, this obviously reduced the percentage of acidic element, with corresponding sacrifice of resistance to thermal shock.

BIBLIOGRAPHY

HODKIN, F. W., and COUSEN, A., *Textbook of Glass Technology*, Constable and Co., Ltd., London (1925).

MOREY, G. W., *Composition of Glass* (lecture), Carnegie Institution of Washington (1936).

MOREY, G. W., *The Properties of Glass*, Monograph Series No. 77, Reinhold (1938).

PITTSBURGH PLATE GLASS CO., *Glass, Paints, Varnishes, and Brushes* (1923).

SCHOLES, S. R., *Modern Glass Practice*, Industrial Publications, Inc., Chicago (1941).

Symposium on Glass, *Ind. Eng. Chem.* **25**, 742, 748, 755, 848, 853, 856, 865, 870 (1933).

TURNER, W. E. S., *The Constitution of Glass*, The Society of Glass Technology (1927).

Chapter XIV

Plasticization by Solution

The fact that many amorphous solids dissolve in suitable liquid solvents to form emulsoidal solutions has already been discussed. If the solid be treated with only a small amount of solvent, the latter is imbibed and the solid swells; it develops a gel structure, its plasticity increases, and its mechanical strength falls. A large ratio of solvent to solid is usually required for complete liquefaction, producing a solution which is correspondingly dilute though frequently very viscous. Evaporation of the solvent will leave the solid behind in its original, relatively non-plastic state.

This method of manipulating plastics has severe limitations. The large amount of solvent required is expensive, except in the case of water. The solid produced by evaporation of the solvent is still sensitive to it and must, therefore, not be used in situations where solid and solvent can come into contact. Where the solvents used are organic liquids, evaporation introduces serious fire hazards. There are often considerable difficulties in removing the solvent. Thus, while the solvent, if volatile, will evaporate, this results in a progressively increasing concentration of the solute. The mass becomes very viscous and ultimately may become practically non-plastic, at the same time shrinking greatly. The solvent can only escape from the interior of the mass by diffusing through it to its surface. As evaporation proceeds, the concentration gradient of solvent, necessary to maintain diffusion, decreases progressively and the diffusivity itself is reduced by the increasing density of the solid, the surface becoming hard and relatively impermeable, so that the solvent can continue to escape only at an extremely low rate. A state is finally reached in which the solvent concentration on the surface is negligible, despite the fact that there may still be

considerable solvent left in the interior. Furthermore, if the solid itself be essentially non-plastic, the final drying may develop serious shrinkage strains. Because of these facts, this method of manipulating plastics is usually limited to the formation of thin films. The manufacture of celluloid and smokeless powder offers exceptions, but when producing such materials in shapes of appreciable thickness the time required for evaporation of the solvent is great, in extreme cases a matter of months. However, for the production of thin films the process has many advantages and is extensively used in adhesives, varnishes, lacquers, photographic films, and the like.

Liquid Adhesives

If one can form between two solid surfaces a thin film of amorphous material characterized by both high internal cohesion and high adhesion to the surfaces, a bond of great strength may be produced. Such a film is often secured advantageously by evaporation of a solution of the adhesive agent. If at least one of the solids being bonded together be permeable to the solvent, the setting action is hastened, though excessive porosity may weaken the bond by carrying the solution away from the surface by capillarity. To secure good adhesion, it is preferable for both the solution and the resultant film to "wet" the surfaces (p. 50). Furthermore, moderately roughened surfaces give anchorage to the adhesive agent. Indeed, the evidence indicates that, in the practical use of adhesives, mechanical embedding of the plastic in the irregularities of the surfaces is of major importance. The volumetric shrinkage will result in a porous and correspondingly weak film unless on drying the bonded surfaces are sufficiently free to move and are maintained under suitable pressure. Such pressure likewise promotes imbedding of the plastic in the surface irregularities.

Because many of the best adhesives are highly emulsoidal colloids, their solutions are likely to be excessively viscous at suitable concentrations. If too dilute, an inadequate film will be left by evaporation unless a large volume of the solution be employed. If too viscous, uniform application is difficult.*

* While in practical use this may cause unsatisfactory bonding, the effect can generally be overcome by adequate care in application of the adhesive and construction of the bond.

The viscosity is often reduced by "breakdown" of the adhesive by suitable treatment, such as the overmilling of rubber to be used in rubber cement or the heating of glue solutions. The effect is presumably due to reduction in molecular weight. The treatment should not be carried beyond the requirements of workability, because the breakdown induces an increase of plasticity of the adhesive which weakens the ultimate bond. Thus, holding a glue solution at boiling temperature reduces its viscosity at first rapidly and then more slowly as time goes on, whereas the tensile strength of the bond produced from it, *e.g.*, between wooden blocks, drops at first slowly and then much more rapidly.

To reduce the expense of non-aqueous solvents, water suspensions or emulsions of water-insoluble plastics are sometimes employed as adhesives. Rubber latex has been thus used, and water emulsions of asphalts are used as road binders. To secure the best bond, such materials require mechanical pressure during drying. The plasticity of the adhesive itself may be enhanced by the presence of a small amount of water-insoluble solvent, preferably one of boiling point sufficiently high so that most of it is retained after the water has been dried out, but low enough to disappear in time.

Gelatin and Glue.

Important by-products of the stockyards and tannery are gelatin and glue, obtained by hydrolytic treatment of bones and skins. Bones contain the protein ossein (or bone collagen), which serves to bond together their inorganic constituents, calcium phosphate and carbonate. Collagen occurs in skins as the main constituent of the derma (p. 379). While insoluble in cold water, collagen slowly dissolves at 60° C., or above, yielding a solution of gelatin which on cooling sets to a characteristic gel. Long pretreatment with either dilute alkali or acid at low temperatures, while not rendering the collagen directly soluble (unless the concentrations are high), induces in it a change which greatly speeds up the rate of solution by hot water, even after removal of the treating agents by washing or neutralization. This may well be due to a breaking of cross-linkages between

protein chains (p. 163). In any event, pretreatment makes possible better yields of gelatin, reducing the excessive hydrolysis which would otherwise accompany the necessarily long heating. The higher the concentration of the alkali or acid, the more rapid the change in the collagen; thus, with dilute caustic soda, a pretreatment of hide trimmings may be effected in twenty-four hours, which would require three or four months with lime at the same temperature. The longer and the more severe the pretreatment of the collagen, the more rapidly it dissolves in hot water but the lower is the viscosity and jelly strength of the product. On prolonged treatment with even dilute alkali sufficient hydrolysis of the protein chains occurs to give detectable quantities of free amino acids.

Gelatin and glue differ essentially in the quality of the raw material from which they are made and in the care observed in their preparation. Edible gelatin is made only from selected, clean materials, such as calves' heads, bones, etc., under hygienic conditions. Photographic gelatin must also be prepared under careful control, avoiding, in particular, contamination by metallic salts. For bulk production, tannery hide trimmings * are an important source. The glue stock, after a preliminary water wash, is treated at normal temperature from four weeks to several months with milk of lime, the mass being turned over from time to time and transferred to fresh lime. The calcium hydroxide causes the protein of the fibers to swell (p. 211), and at the same time loosens or dissolves objectionable material such as hair, blood, and the easily soluble proteins (mucins, globulins, and albumins), which would contaminate the gelatin. Fatty impurities are converted to insoluble lime soaps, which rise to the surface as scum, removable by washing through a screen to retain the stock. While lime, because of its limited solubility, is slow in action, it automatically holds the pH almost constant and excess is readily washed away in suspension after treatment, whereas removal of soluble alkali, such as caustic soda, is difficult. During washing, a considerable amount of stock, dependent upon the time of pretreatment, is lost by solution. The washed

---

* Chrome-tanned leather scrap may be detanned with oxalic acid and subsequently treated like other stock, but its quality deteriorates under the treatment.

stock is neutralized, often with sulfurous acid,* which also serves as a bleaching agent and an antiseptic against bacterial action. In this way, the lime is almost completely removed. The washed stock is now extracted repeatedly with water. The first extraction, made at 60 to 70° C., removes the most soluble material with little hydrolysis of the extracted gelatin. There follow four or five extractions at successively higher temperatures, the last sometimes made above 100° C. in an autoclave. To improve uniformity of product, corresponding extracts from a number of batches are mixed. First extract products have excellent gelling power, while later extracts give inferior grades of glue. The solutions are clarified, usually by filtration. Very clear, sparkling gelatins are obtained by adding egg or blood albumin to the solution, coagulating by slowly raising the temperature, and siphoning off the clear liquid. The solutions set on cooling to a firm jelly of high water content. By means of wires stretched on a frame the jelly is cut into slices, and these dried on trays † to a hard, bony sheet. Because of the low liquefying point of the jelly, the temperature of drying must not be above 40° C., and yet, because of tendency to bacterial decomposition, the drying operation must be as rapid as possible.

Bones are first degreased with organic solvents, and calcium phosphate and other mineral constituents are dissolved by cold, dilute hydrochloric acid. Thereafter, the procedure is essentially that outlined above. Fish glue is made from fish residues; because of its low jellying power, it appears on the market as liquid glue. Even concentrated solutions of the best glues are fluid at temperatures above about 35° C. For use as adhesives, they should always be applied in this condition. Because this involves the inconvenience of maintaining the glue pot at elevated temperatures, so-called liquid glues are often produced by treatment of ordinary glue with acid, usually acetic. These give solutions of high concentration, which remain fluid at ordinary temperatures. Presumably, presence of acid decreases the gelation tendency by increasing solubility (p. 211).

* Except in the case of edible gelatin, where the amount allowable in the final product is frequently limited by law. Phosphoric acid is commonly used in this case.
† Spray driers have been used for both gelatin and glue.

Vegetable Adhesives.

Starch is obtained from certain vegetable tubers and seeds, e.g., potatoes, corn, rice, etc., by steeping and mechanical grinding in water. Starch granules are set free and can be separated from the cellulose, protein, oils, and other impurities by continued washing. The water in the resulting cream may be removed by filtration, often on rotary filters, and the starch dried on trays in an atmosphere of controlled temperature and humidity.

When heated with water, the cell wall surrounding the individual starch grains breaks and a thick, viscous paste results. The temperature of rupture, while varying from 55 to 85° C. for different starches, is nearly constant for a given species. Starch in solution is not homogeneous, since it is possible to separate it, e.g., by centrifuging under proper conditions, into α-amylose, insoluble in cold water and giving a purple color with iodine, and β-amylose, which yields a clear blue color with iodine and, freshly prepared, is soluble in cold water. Both contain a small proportion of phosphorus, perhaps as ester. Both consist essentially of α-glucopyranose units (p. 152) linked in the 1–4 positions:
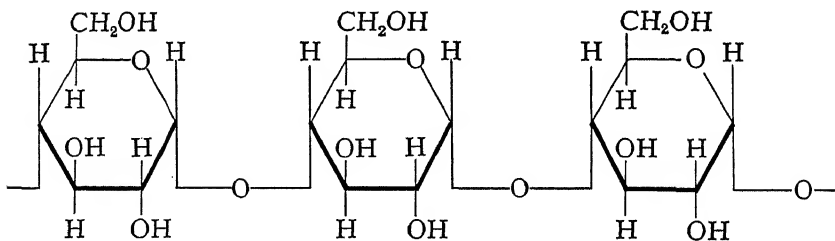
The reason for the differences is not certain, but may be a matter of chain-length.

A starch solution has too great a viscosity, even at low concentrations, for successful use as an adhesive. However, as a polysaccharide, $(C_6H_{10}O_5)_x$, it can be hydrolyzed in successive stages to give products more and more soluble in water, with progressively lower viscosities at identical concentrations. The final product, using dilute acid as the hydrolyzing agent, is glucose (p. 151). Among the intermediate products are found

a number of important adhesives. In industrial practice there are three principal methods for controlling the hydrolysis:

1. Heating the solid starch after moistening with dilute acid or other hydrolyzing agent.
2. Heating in solution with dilute acid.
3. Heating in solution with dilute alkali.

Dextrins, irrespective of the amount of breakdown, are products of the first two processes. British Gums are prepared by heating with little or no hydrolyzing agent at a somewhat higher temperature (300 to 400° F.) than is usual in the case of dextrins (230 to 300° F.). All these products vary in viscosity, color, adhesive power, etc., with the type of starch used and the severity of the hydrolytic treatment employed.* Starches are often blended. Potato starch gives a dextrin with good adhesive qualities, but objectionable in taste and odor. Tapioca starch, which does not suffer from these disadvantages, is generally used for the preparation of gums for envelopes, postage stamps, and similar materials. Dextrins from corn starch are used for many lines of work and probably represent the greatest tonnage. Where the temperature is kept down, the hydrolysis can be carried out in water suspensions without rupturing the granules. This is done in making the so-called thin boiling starches, prepared by digesting a starch cream at 100 to 125° F. with dilute hydrochloric or sulfuric acid. Hydrolysis occurs within the granule, so that when it bursts in boiling water a relatively thin solution is produced. These products are particularly useful for sizing and stiffening cotton fabrics, since on drying down they give a film, which, in contrast to that produced by dextrin, is not brittle and resists solution in water.

The second method, partial hydrolysis in solution, makes it possible to minimize breakdown and produce adhesives of high bonding power. However, the process suffers from the disadvantage that the product must be prepared at the point of use or else recovered from the solution.

* Hydrolyzing agents commonly used are hydrochloric, nitric, sulfuric, and oxalic acids. Volatile acids are to be preferred for the first method, perhaps because of better diffusion through the starch granules, while non-volatile acids are generally used in the second.

The third process yields a product known as vegetable glue. Starch is suspended in about 3 parts of water and 5% caustic soda solution is added. On heating until the starch granules burst a glue-like mass of good bonding power, capable of replacing animal glue in joining woodwork, etc., is obtained.

Besides glues obtained from starch, certain natural gums, such as tragacanth and arabic, are sometimes employed. Solutions of rubber and of nitrocellulose in organic solvents also yield adhesives of moderate strength. In recent years the latter type has become increasingly important as cement in the manufacture of shoes.

PROTECTIVE COATINGS

It has long been common practice to protect structural surfaces against both mechanical and chemical action (e.g., wood against weathering, metals against corrosion, etc.) by a thin film of a suitable amorphous material. In this way a hard, smooth, attractive finish, highly resistant to abrasion, atmospheric attack, and the like, can be produced. By incorporating suitable pigments, it is possible to secure increased protection and a wide variety of attractive color effects. A helpful classification is given in Table I. A varnish is fundamentally a solution in a volatile solvent of a film-forming, inert, resinous amorphous base, non-plastic as such. Because the resins available are frequently too brittle for satisfactory use, they are often softened by the incorporation in the solution of a certain amount of drying oil, giving the oil varnishes in contradistinction to the spirit varnishes. Where pigment is incorporated in the varnish, the product is usually described as an enamel. In recent decades there has come into use an important modification, known as lacquers, in which the essential amorphous base is a cellulose ester, rather than the resins characteristic of the varnishes. They may or may not be pigmented.

Paints are fluid mixtures, consisting of pigments suspended in a liquid, which can be applied as thin films to surfaces by brushing, spraying, or equivalent technique, and which will then set spontaneously to an inert, protective coat. Ordinary paints consist essentially of finely ground opaque or semi-opaque

inert substances, known as pigments, dispersed in a drying oil, called the vehicle. On drying, this furnishes the film which bonds the pigment together and to the surface. Suitable solvents or thinners may be added to obtain correct consistency for application; to accelerate drying, small amounts of metallic

TABLE I.

PROTECTIVE COATINGS

Volatile solvent as
major plasticizing
agent

Oil vehicle —
deplasticization
mainly by oxidation

ORDINARY PAINTS

Cellulose ester
base

Colloidal binder —
water solvent

LACQUERS

CALCIMINES

[with or without
pigments]

Resin Base

Pigmented

Unpigmented

ENAMELS

Volatile solvent only

Drying oil present

SPIRIT VARNISHES

OIL VARNISHES
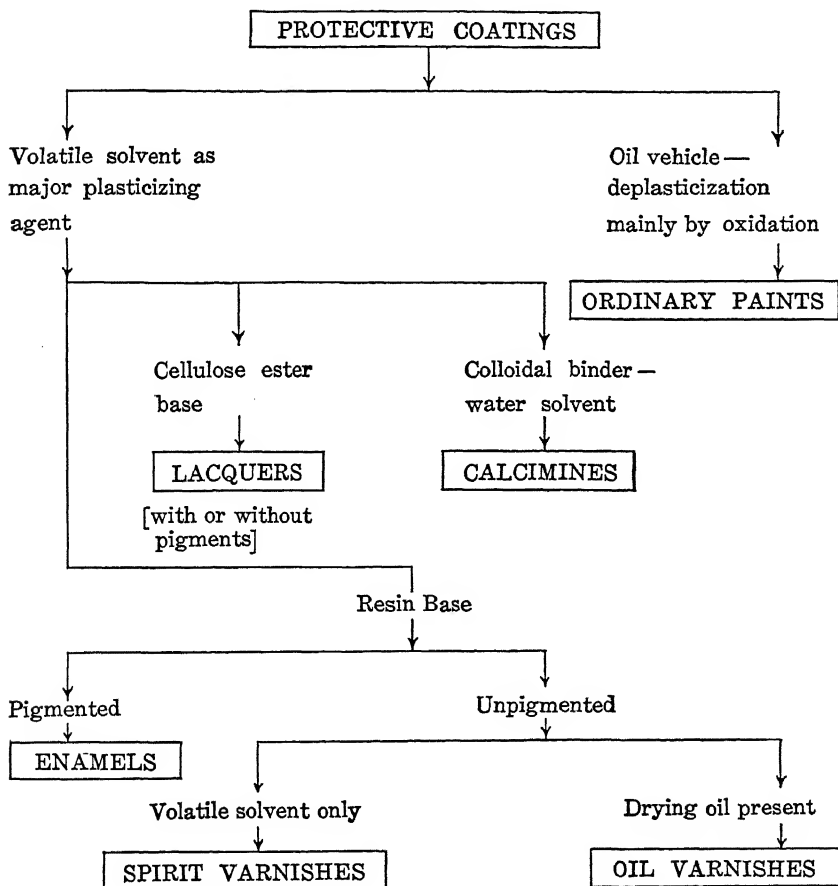
salts, catalytic in effect, commonly known as driers, are incorporated with the vehicle. Paints may be made up with water as the liquid medium, using suitable water-soluble colloids as bonding agents, such as dextrin, casein, glue, etc. Distempers or calcimines belong in this category.

Varnishes.

The resins, which constitute the base of varnish films, are solid, amorphous materials, more or less colored, vitreous in structure and consequently capable of forming surfaces of high luster. They usually melt at moderate temperatures, often, however, with considerable thermal decomposition. Formerly, the most important of the varnish resins were natural products, but of recent years synthetic resins (p. 469) have come into use and because of uniformity of quality and moderate cost have become increasingly popular. The natural resins are usually tree exudations which on exposure to the atmosphere dry into hard, semitransparent masses. Such masses may be collected almost as soon as formed,* or may, by chance, have been buried in the ground over considerable periods of time and subsequently dug up, in which case they are known as fossil resins. Their most important characteristics are hardness and solubility, a continuous series of gums being available from the hard fossil copals to the soft, freshly tapped dammars. The resins are related chemically to the terpenes and essential oils. As indicated by their acid numbers, which vary from about 20 to 150, they always contain considerable quantities of acid bodies. Their exact composition is, however, unknown.

The softer varnish resins, dammars and mastics, are soluble in alcohol and similar organic liquids, the so-called spirit varnishes being essentially such solutions. One of the most important of these is shellac,† which, dissolved in alcohol, gives a quick-drying, lustrous finish. Spirit varnishes give hard films of considerable brilliance, but in time these may develop fine cracks or fissures which disfigure the surface.‡

The harder, fossil resins are practically insoluble, but when heated to temperatures in the neighborhood of 300 to 350° C. undergo a thermal decomposition with evolution of vapor con-

* The major sources of the natural varnish gums, which are frequently named for their point of origin, are the East Indies, New Zealand (Kauri), and Africa.

† Lac is the excretion induced by an Indian insect, which during the major portion of its life-cycle attaches itself to a tree twig. The soft, resinous exudation subsequently hardens and with those formed by thousands of other similar insects completely covers the twig. After detaching and undergoing minor purification, it yields shellac.

‡ Similar materials are the asphalt paints, which are solutions of asphalts and related materials in volatile solvents, preferably aromatic in character. These are also sometimes known as lacquers.

stituents amounting to as much as 25% by weight of the gum, and become soluble. It seems probable that the change in properties is due to a thermal disintegration or depolymerization.* While the resin is now soluble it is rarely if ever used alone, but is mixed with a drying oil, previously heated and correspondingly polymerized (p. 324), so that the final film sets not only by evaporation of the solvent employed but also by oxidation of the oil.

"Gum running" is carried out in relatively small, direct-fired kettles of copper, aluminum, or stainless steel, equipped with covers vented to flues for elimination of the acrid, highly inflammable vapors. The resin is introduced and the temperature brought up slowly.† When the decomposition has proceeded to the desired point, the hot linseed or tung oil is introduced and the mixture is held at a high temperature level to secure complete homogeneity. After cooling to the neighborhood of 200° C., it is diluted to proper consistency with turpentine or a carefully refined petroleum cut of suitable solvency and volatility. The cold varnish is clarified by separating imperfectly dissolved fractions. This was formerly done by gravity settling in large vats and sometimes required many months. Filtration might also be employed. Centrifugal clarification is now widely used, but it seems probable that some beneficial change other than sedimentation occurs on ageing; storage is still used for certain high-grade products.

The initial drying of a varnish film is due to evaporation of the solvent, the rate being determined mainly by the volatility. The final hardening of the film is relatively slow because it is controlled by the oxidation of the tung oil fractions. The larger the ratio of oil to resin, the softer the film after solvent evaporation and the tougher and more flexible after final drying. Both stages of drying may be accelerated by heating (stoving) at moderate temperatures (150° C.), a process which, properly controlled, also yields a denser and more durable film. The use of these so-called baking japans is obviously restricted to ob-

---

* Mechanical manipulation, similar to the milling of rubber (p. 411), is capable of effecting a similar change at far lower temperatures.

† A small amount of the previous batch may be left in the kettle to serve as a menstruum for dissolving the resin, thus reducing localized overheating.

jects, usually metallic, of limited size.*   Finishes of this general type are the basis of so-called patent leather (see pp. 394–396).

For high-grade work, multiple coats of varnish are always used.   The individual coats can thus be very thin, facilitating complete evaporation of solvent and avoiding a relatively hard surface layer over a more plastic interior.   To secure uniformity of the final film and to improve bonding between coats, the dried surface is sandpapered between successive applications.   In finishing wood, this operation opens the pores of the surface and allows penetration of the newly applied varnish into the remaining voids of the wood itself, thus developing a density and consequent durability of surface otherwise unattainable.   By modification of composition of the successive coats, both the toughness of the total film and bonding power between the hard external finish and the protected surface can be improved.

Synthetic Varnish Resins.

The natural resins suffer from the fact that in order to be soluble they must not have too complex a molecular structure, although complexity promotes film hardness and resistivity. The gum-running operation, therefore, represents a sacrifice in quality to get workability.   Moreover, once the resin is depolymerized, no methods are available for restoring its original structure by repolymerization.   In sharp contrast to the natural resins, the synthetic resins (p. 469) require no thermal breakdown (with its tendency to discoloration and formation of undesirable by-products), but can be applied in solution at a relatively early stage of polymerization or condensation, and after the drying of the film polymerization to any desired degree can be secured by a subsequent operation such as baking. Furthermore, at the low stage of polymerization at which they are used in the varnish, they have maximum compatibility with the oil.   This gives ease of manufacture and high quality of product which, together with adequacy of supply, is resulting in rapid extension of the use of synthetic resins in the varnish industry.

---

* For black finishes the resin is frequently replaced by asphalt, with suitable modification of the solvent.   Somewhat higher baking temperatures may be employed in this case.

The synthetic-resin varnishes are characterized in general by high resistance to weathering and chemical action, elasticity of film (requiring lower ratios of oil to resin to secure adequate flexibility), good adhesion (particularly to metal surfaces), high luster, and, in some cases, high color stability to both light and heat. It is frequently possible to secure an improved combination of desirable characteristics by suitable blending of resins of different types.

The phenol-formaldehyde resins (p. 473) give hard, brilliant films, highly resistant, except for low color stability, to both light and heat. They are frequently modified by incorporation of rosin or rosin esters, reducing cost without great sacrifice in quality. For use in varnishes the alkyd resins, e.g., those based on glycerin and phthalic anhydride (p. 155), have fatty acids, preferably unsaturated, incorporated in the original mix, increasing the solubility of the resins and flexibility of the ultimate film (p. 481). The fact that the fatty acids are chemically combined in the ultimate molecular structure probably contributes significantly to the resistance of the films and their combination of flexibility and luster. The condensation products of maleic anhydride with rosin and its esters (diene reaction) are flexible, but have hitherto been expensive. They are slow in drying and the films lack the resistance to water and chemicals of the phenolic-aldehyde resins, but exhibit excellent color stability to sunlight and heat (including baking). They are consequently useful for white enamels and light tints. Urea-formaldehyde resins (p. 477) are rapidly coming into extensive use for baking enamels because of their extraordinarily rapid polymerization under heat and the outstanding color stability and hardness of the resultant film. They are applied in solution, e.g., in butanol and coal tar solvents, and baked a few minutes in the neighborhood of 175° C. The brittleness of the film is usually corrected by blending, e.g., with alkyd resins. They are ideal for white enamels. Unfortunately they do not air-dry properly.

Lacquers.

The attempts to improve quality and reduce cost of varnish films have resulted in the use, not only of the synthetic resins,

but particularly of nitrocellulose and other cellulose derivatives. This development was stimulated after the War by the availability, first, of large stocks of nitrocellulose manufactured for smokeless powder, and second, of more adequate sources of suitable organic solvents.

Nitrocellulose had long been used as a protective film, but because of viscosity could be applied only in very dilute solutions. Furthermore, the film had poor adhesion, particularly to metals, was a poor bond for pigments, was relatively brittle, inferior in resistance to weathering, and underwent serious chemical degeneration with age, particularly on exposure to light. These disadvantages have been overcome in recent years by an important series of developments.

The characteristics of emulsoidal solutions (p. 164) and the structure of the cellulose molecule (p. 151) make it probable that the extremely high viscosity of the solutions of cellulose esters is due to chain-like molecules of high molecular weight. In view of the known degradation of cellulose by hydrolysis, the possibility suggests itself of thus reducing the viscosity of the ester solutions, either before or after esterification. It was discovered that by heating the normal nitrate at temperature levels of approximately 130° C., under water held liquid by pressure, it undergoes a change, involving no substantial reduction in nitrogen content (which is approximately 11 to 12%), but accompanied by great reduction in viscosity of its solutions. This was the first essential step in rendering practicable the widespread use of cellulose lacquers. However, this results in an ultimate deterioration in film quality, so that it is undesirable to carry the transformation further than necessary. By the addition to such a solution of a liquid miscible with the solvent but not itself a solvent for nitrocellulose, still further reduction in viscosity can be achieved. It seems probable that the explanation is a desolvating effect on the cellulose molecule, similar to that discussed on p. 175. The combination of these two techniques makes it possible to produce cellulose ester solutions of satisfactorily high concentrations and yet of workable viscosities.

Because no film-forming constituents have ever been found superior to the varnish resins in gloss and in resistivity to physical

and chemical disintegration, these are incorporated in the lacquer. They reduce permeability to water and increase adhesion to metallic surfaces. In effect, this use of gums amounts to replacing the drying oil of an oil varnish by the cellulose ester, thus taking advantage of the superior film-forming characteristics of the latter, and also of the possibility of extremely quick drying, since no oxidation or other chemical change subsequent to evaporation of the solvent is necessary. On the other hand, nitrocellulose does not have the softening effect on the film which the drying oils possess; to overcome this deficiency it has been found necessary to incorporate a group of materials, the so-called plasticizers. These are high-boiling, non-volatile liquids, miscible with the cellulose nitrate and the gums employed. They not only reduce the brittleness of the film but improve its adhesion to metals and its bonding power for pigments, probably by increasing the flexibility of the residual film. Chemical degeneration of the nitrocellulose is greatly reduced by incorporating as inhibitors traces of organic bases such as urea, β-naphthylamine, and the like.

These changes opened, however, a Pandora's box of new problems, mainly connected with the solvents. The solubility relations of a given nitrocellulose depend both upon its nitrogen content and the degree of degeneration to which the cellulose or the ester has been subjected. The behavior toward solvents is abnormal. Thus, a nitrate suitable for lacquers is insoluble in either ethyl alcohol or ethyl ether but soluble in a mixture of the two.* The effective solvents are of intermediate polarity (p. 15), although polarity is certainly not the only factor involved. The esters of fatty acids and aliphatic alcohols, both of low molecular weight, have high solvent power, e.g., ethyl, butyl, and amyl acetates; the ketone group comes next in solvent action. The simple ethers are not solvents but become so if sufficient ether oxygen be introduced into the molecule; thus, glycol diethyl ether is better than ether, and diethyl carbitol, C_2H_5—O—C_2H_4—O—C_2H_4—OC_2H_5, is better still. The best solvents for nitrocellulose are not the best for the

---

* There is a general tendency for an alcohol group to increase the solvent capacity; e.g., ethyl lactate, and diacetone alcohol (see p. 370).

varnish gums. In consequence, mixed solvents are always‚used and the relationships become extraordinarily complicated. Thus, advantage can be taken of the fact that, while alcohols alone are not nitrocellulose solvents, they can not only enhance the effectiveness of true nitrocellulose solvents but are also among the best solvents for the resins. Hydrocarbons, particularly of the aromatic type, while non-solvents of nitrocellulose, have some solvent power for many of the varnish resins and can be incorporated in the lacquer to a considerable degree without precipitation of the nitrocellulose. They have the advantage of markedly reducing the cost. All of these materials which are used primarily to reduce the viscosity of the lacquer to the point essential for proper application are called diluents, although the term is sometimes reserved for the hydrocarbons. Because the total quantity of solvent in the lacquer is dictated mainly by the fluidity required, it is often possible to reduce cost by the use of a nitrocellulose solvent, which, though expensive, has high tolerance for the diluent. The adjustment of these solvent requirements with the necessary evaporation characteristics of the mixture constitutes the problem of solvent balance. One must employ solvent-diluent mixtures which will not only give initial homogeneity but maintain it in the final film, and preferably throughout the period of drying. If at any time the proportion of diluent to solvent be excessive, nitrocellulose will separate out to a greater or less degree. If the ratio of solvent to diluent be excessive, separation of gum will occur. If either form of non-homogeneity develops in the later stages of evaporation, the plasticity of the film is so low that it persists in the ultimate film, giving it a chalky opalescence. Where this is induced by separation of nitrocellulose, the effect is called cotton blush; the corresponding phenomenon due to separation of resin is gum blush. However, incipient separation of either of these constituents in the early stages of drying may disappear if the solvent-diluent ratio later rectifies itself.

The ideal situation would be the use of a solvent-diluent mixture which maintains equal (though not constant) volatilities of the constituents throughout the process of drying, so that the mixture does not change its composition. This is practically

impossible to achieve, particularly because the mixtures employed almost always show abnormal vapor pressures. Furthermore, the conditions of drying affect the relative rates of evaporation.* Changes in the solvent-diluent ratio are therefore sure to occur during drying, but, to avoid undesirable results, must be controlled by proper composition of the mixture of volatile constituents.

To secure the potential advantages of rapid rate of drying, one is tempted to use solvents of very low boiling point. This can result in another type of blemish, so-called water blush, caused by condensation of moisture on the surface of the film from the water vapor in the air. In the case of a solvent completely immiscible with water, this can occur only if the surface be cooled by evaporation of the solvent † to a temperature below the dew point of the water vapor in the air. However, a solvent miscible with water can dissolve considerable amounts from the air and still have the partial pressure of water vapor over the solution far lower than that of pure water at the temperature of the evaporating surface.‡ Owing to the reduction in solvent power, induced by the dissolved water, this can cause water blush at relatively high surface temperatures. The general interrelations of these effects and

		SLOW DRYING		
GUM BLUSH	Permanent	Excess High Boiling Solvent		Excess High Boiling Diluent	Permanent	COTTON BLUSH
			Correct Composition (Solvent Balance)		
	Temporary	Excess Low Boiling Solvent		Excess Low Boiling Diluent	Temporary
		WATER BLUSH		

FIG. 1. Solvent Balance. Effect of the relative proportions of the volatile constituents on the film.

the factors entering into the composition of ester lacquers are summarized in Fig. 1 and Table II, respectively.

* Drying by evaporation in the presence of an inert gas, such as air, does not completely parallel ordinary distillation. The phenomenon is complicated by the diffusion of the vapor constituents from the surface of evaporation out into the body of gas, components of low molecular weight traveling more rapidly and hence possessing abnormally high volatilities.

† This is equivalent to the wet bulb depression of a thermometer reading, as employed in measuring air humidity.

‡ Fortunately, many solvents and solvent mixtures form azeotropic mixtures of minimum boiling point with water, and hence the volatility is greater than that of the anhydrous solvent. Provided the water absorption does not exceed the azeotropic mixture, this results in preferential elimination of the water on evaporation, and corresponding reduction in the risk of ultimate water blush.

Other cellulose compounds, particularly the acetate and various of its ethers, are available for lacquer use. They have the advantage of greater resistance to deterioration, particularly

TABLE II. COMPOSITION OF CELLULOSE ESTER LACQUERS

Component	Function	Example
1. Base (*a*) Cellulose derivative (*b*) Gums	To give strength and toughness to film To give gloss, adhesive power, and impermeability	Cellulose nitrate, acetate, etc. Ester gum,[a] dammar, copals, etc.
2. Solvents (*a*) Low boiling (*b*) High boiling	To dissolve cellulose ester (sometimes gums also) For rapid drying For slow drying	Esters, ketones, ethers, volatility depending upon molecular structure
3. Diluents (*a*) Low boiling (*b*) High boiling	To reduce cost To dissolve gums For rapid drying For slow drying	Hydrocarbons, alcohols, volatility depending upon molecular structure
4. Plasticizers	To give flexibility	Dibutyl phthalate, tricresyl phosphate, castor oil, etc.
5. Stabilizers	To prevent deterioration (decomposition of the nitrate)	Urea, β-naphthylamine, etc.
6. Pigments—dyes	To color and protect base against ultraviolet light	Inert inorganic compounds and organic dyes

[a] This is made by heating together rosin (largely abietic anhydride) and glycerol, leading to a highly polymerized ester (cf. pp. 140–145, 155–157).

on exposure to light and elevated temperatures. However, the acetate, one of the most readily available, is difficult to get into clear homogeneous solution when mixed with the resin required. The esters of higher fatty acids are more compatible with the varnish gums; thus, good results are obtained by using a mixed ester containing roughly 20% butyrate and 40% acetate radical (p. 377). Ethyl cellulose (p. 378) is also used to some extent.

Lacquer Emulsions.

One of the limitations of lacquers is the expense of the organic solvents required for their solutions. A recent development is

the introduction of nitrocellulose lacquers plasticized by a relatively small amount of a water-immiscible solvent mixture of the type described above, yielding a mass far too viscous for direct application. This material is thinned to suitable workability by dispersion as a water emulsion. Materials compatible with the lacquer film on drying should be used as emulsifying agents, mixtures of methyl cellulose with sulfonated castor oil or similar compounds being particularly satisfactory. Emulsification cheapens the lacquer and greatly reduces the fire hazards involved in its use. Perhaps even more important is the possibility of using a nitrocellulose which has suffered less degradation, yielding correspondingly stronger films.

Paints: Vehicles.

The vehicle of a paint is the liquid medium in which the solid pigments are dispersed. The major constituents are the drying oils—linseed, china wood or tung, and, less frequently, perilla.* These oils consist predominantly of glycerol esters of fatty acids, most of which are highly unsaturated. Thus, the acids of linseed oil are approximately 18% oleic ($C_{17}H_{33}COOH$, with one double bond in the hydrocarbon chain), 37% linoleic ($C_{17}H_{31}COOH$, with two double bonds in the chain), and 40% linolenic ($C_{17}H_{29}COOH$, with three double bonds). The double bonds give the oil an iodine number of 160 to 190, depending on the source of the oil and method employed in testing. It always contains some free acid, with acid values in the neighborhood of 6. The importance of the drying oils lies in the fact that in a thin layer they absorb up to about 15 to 20% of their own weight of oxygen from the air, and are thereby transformed into an elastic, insoluble film, which in the case of linseed oil is called linoxyn. Such films adhere firmly to wood and most metals. However, they are permeable to water, have low resistance to abrasion, and deteriorate seriously on exposure to the weather. On the other hand, when loaded with pigment particles, for which they serve as effective binders, their physical properties

---

* Chemically modified oils are becoming competitive. Thus, it is possible to eliminate water from the hydroxyl group of the fatty acid chain of ricinoleic acid in castor oil, $CH_3(CH_2)_5CHOHCH_2CH=CH(CH_2)_7COOH$, forming the active conjugated diolefinic linkage, $CH_3(CH_2)_5—CH=CH—CH=CH(CH_2)_7COOH$, which gives drying properties to the oil.

and chemical resistance are improved to such an extent that these films form a most important type of protective coatings for wooden and metallic surfaces.

It is a surprising fact that the mechanism of "drying" is unknown. In the initial stages, as already stated, large quantities of oxygen are absorbed, much of it forming peroxides. However, the oil remains fluid until a later stage in the process, during which further oxygen absorption is relatively slight. The molecular weight of the dried film, cryoscopically determined, is over twofold that of the original oil, but there is the possibility that changes in molecular association occur in bringing the film into solution. During drying, the iodine number decreases, indicating elimination of double bonds. If the oil is held at high temperatures in the absence of oxygen, polymerization occurs, indicated by rise in molecular weight, decrease in iodine number, increase in viscosity, and ultimate setting to a semi-solid gel. These changes parallel those of oxygen drying. Indeed, it is common practice thus to treat drying oils for use in paint and varnish vehicles, to cut down the time necessary for air drying, yielding boiled or stand oils. By heating at somewhat lower temperatures, the oil can be transformed by blowing air through it, producing blown oil. The parallelism in changes of properties of the oil induced by heating alone and by oxidation leads to the assumption that the changes are due to fundamentally similar chemical transformations. Since the increase in molecular weight and decrease in unsaturation induced by heating alone point to a polymerization due to cross-bridging between molecules at the olefinic linkages, one assumes a similar mechanism for the actual drying by oxidation. The former almost certainly involves carbon-to-carbon linkages between the molecules, but whether the intermolecular bridging due to oxidation involves oxygen linkages is unknown.

One difficulty in arriving at a clear understanding of the oxidation reaction lies in the fact that, while oxygen tends to attack all olefinic linkages, such attack seldom results in the development of intermolecular linkages between the olefins concerned. The formation of peroxides by the addition of oxygen to olefins is common. On the other hand, such com-

pounds usually decompose with rupture of the carbon chain at the olefinic bond. This is perhaps the source of the carbon dioxide and volatile fatty acids and aldehydes of low molecular weight, which are formed in the drying of oils. Because the percentage of these is small, it seems unlikely that their formation is a major factor in the drying process itself. Apparently drying is due to breakdown of peroxide linkages formed initially on olefinic double bonds, the mechanism of breakdown resulting in cross-bridging between molecules rather than rupture of hydrocarbon chains at the olefinic bonds.

Olefinic compounds vary greatly in susceptibility to oxygen addition. The inactivity of single double bonds in a fatty acid chain, as in oleic acid, is evidently the reason for the resistance of triolein to drying. The conjugated diolefinic bonds of linoleic and linolenic acids not only are far more active in oxygen addition, but also provide a much larger number of active points in the molecular structure between which intermolecular cross-linkages can develop (see p. 155). In keeping with this, neither these acids nor their monoesters, although oxidizing readily, will dry satisfactorily, as do their triglycerides. In the latter case one has the advantage of starting with a material of approximately three times the molecular weight and threefold possibilities for cross-bridging.

Specific catalysts of the drying reactions have long been recognized. These driers, in order of increasing effectiveness, are the lead, manganese, and cobalt salts of organic acids, *e.g.*, the fatty and naphthenic acids. The metals which are effective are all readily capable of existence in more than one stage of oxidation and presumably operate as oxygen carriers. The salts employed are oil-soluble,* and about 0.1% of the weight of the oil is added. It seems not unlikely that the drier affects mainly the decomposition reactions of the peroxides initially formed. This is indicated by the fact that, at a given stage of drying, the oxygen content of oils containing drier is lower than otherwise. Furthermore, the use of driers increases somewhat the percentage of volatile products. The drying reactions are likewise

---

* While one can use the oxides of the metals, these are apparently effective only as they react with fatty acids in the oil itself.

catalyzed by light, particularly by ultraviolet, making it possible to get a harder film for a given oxygen absorption. Thus, for years it has been common practice to get the final finish on the linseed oil coatings of patent leather (p. 395) by exposure to sunlight.

The mechanism of initial oxygen addition is autocatalytic in the sense that rate increases with time. The induction period can be greatly increased by compounds of the type of antioxidants used in rubber, cracked gasoline, etc., *e.g.*, oxidizable phenols, amines, and the like.

Pigments.

The pigments profoundly influence the character of the final film. Thus, the introduction of driers was originally made necessary by the substitution of zinc oxide for white lead. The latter reacts with the fatty acids of the oil to a degree sufficient to give the catalysts essential for satisfactory drying, while the former, although also reacting, has no significant drying effect. Other influences of the pigments are apparently due mainly to physical characteristics. Thus, a mixture of high-grade pigments, such as white lead and zinc oxide, will give a more satisfactory film than either alone. The properties of such a film can be still further improved by incorporating a certain percentage of far coarser pigments, which by themselves are practically worthless—the so-called extenders. This calls to mind the well-recognized principle of mixing in the case of materials such as concrete, where a mixture containing coarse, skeletal particles has the voids between these filled with progressively finer particles, and the whole mass ultimately bonded together by a suitable matrix. The presence of 10 to 20% of extenders in a paint, rather than being an indication of adulteration, may add materially to quality. Particle shape of the extenders is important. Thus, platy, micaceous extenders can reduce the tendency to settling in the mixed paint by the development of a structure in the mass, while in the final film the plates will be smoothed down by the brushing operation into positions parallel to the surface, giving a scale-like, shingle effect which can add materially to the impermeability of the film.

Flow Characteristics of Paints.

The consistency of a paint is an important characteristic. It must flow readily and smoothly under the action of the brush; after application, it must flatten out under the surface tension of its vehicle, so as to eliminate brush marks and other irregularities; it must not run down on inclined surfaces nor sag off the underside of horizontal ones. This combination of properties can only be secured in paints acting as non-Newtonian liquids (see p. 137). It must have a yield point high enough to prevent running or sagging. Furthermore, it should be thixotropic, readily becoming fluid under the action of the brush, but setting after brushing ceases. However, the thixotropic rate of setting must be slow enough so that brush marks will eliminate themselves before gel formation develops. This non-Newtonian flow and thixotropic action is achieved by using a pigment concentration sufficiently high to give gel-like characteristics to the fluid,* as in the suspensions of clays, quicksand, and the like (p. 243). Particle shape has an important effect on flow characteristics, needle-shaped and platy particles giving yield points at relatively
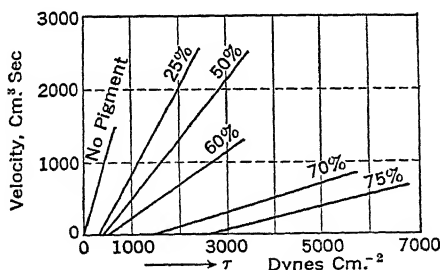
FIG. 2.   Consistency Curves of a Lithopone-Linseed Oil Paint.†

low concentrations. Thixotropic characteristics have the secondary but by no means unimportant advantage of reducing the settling of the pigment particles in a paint on storage. A series of typical viscosity curves made with a capillary viscosimeter, for a lithopone-linseed oil paint, are shown in Fig. 2. The oil

* A small percentage of liquid water is sometimes dispersed in paint to aid in developing this thixotropic effect.
† HOUWINK, R., *Elasticity, Plasticity and Structure of Matter*, Cambridge University Press, Macmillan (1937).

behaves as a Newtonian liquid, with its curves a straight line through the origin; as lithopone is added, yield point develops and mobility decreases.

Paint Mixing.

The manufacture of a satisfactory paint requires the complete dispersion of the pigment in the vehicle. This is accomplished by grinding in oil, a process which, however, does not subdivide the ultimate particles of the pigment to any significant extent, but does break up pigment flocculates into the individual particles. Pigment is first incorporated in the oil in powerful mechanical mixers and the paste is then ground. Grinding was formerly conducted almost exclusively in buhrstone mills, but modern technique favors pebble and roller mills.* The stone and pebble mills are conventional in type,† but the roller mills are almost specific to the industry. To secure capacity the number of rolls has steadily increased, Fig. 3 showing diagrammatically a five-roll mill. The feed is delivered to the two bottom rolls and is picked up progressively by the three rolls at the left located in a vertical plane, each of these running at higher speed than its predecessor, developing both pressure and shear. Roll mills are scratched by earthy pigments, but otherwise have wide utility.

FIG. 3.   Five-Roll
Paint Mill.

Heat generation is a serious problem in paint grinding. Stone and roller mills are both water-cooled, the water chamber in the former extending over the whole back surface of the stones.

* The stone mills, of buhrstone, black rock, or gropus, consist of two flat, superimposed circular discs about 4 ins. thick. The contacting faces of both stones are furrowed from the center to about 3 ins. from the periphery. The furrows are about 3 ins. apart at their radial ends and vary in depth from about 1/2 in. near the center of the stone to a feathered edge at the end of the furrow. Grooves lead from the furrows at approximately 20° angles and are continued to about 1/4 in. from the edge, feathered to this point for fine grinds. For ordinary paint grinds, where production is more important than maximum dispersion of pigment, the grooves lead to the edges of the stones. The finer the furrows and grooves, the longer it takes for the mix to pass between the stones and the better the dispersion of the pigment particles in the grinding medium. The bottom stone revolves about 80 R.P.M. Production of pastes for fine enamels on a 26-in. diameter stone mill is about 15 lbs. per hour, while for ordinary paints the production runs up to five times that amount.

† Where scale of operation is limited the industry occasionally uses very small mills.

The most serious limitation of pebble mills is the difficulty of cooling them effectively, but a further difficulty is that, unless lined with stone or porcelain and using flint pebbles or porcelain balls, they are unsuited for light colors because of contamination.

In all cases the action is predominantly one of shear, well calculated to separate the particles into a perfect dispersion. There is an optimum oil-pigment ratio for grinding, involving the presence of enough oil to make complete particle dispersion possible, but otherwise keeping pigment concentration high to accentuate inner particle friction as the mass slips between the rolls. This optimum oil-pigment grinding ratio is usually too low to give an ultimate paint film of maximum strength and resistivity. In consequence, the paste leaving the paint mill is thinned with raw or boiled oil to the proper proportion for the finished film. This in turn frequently gives a paste too thick for proper consistency of the finished paint, so that it is further thinned to working consistency by volatile solvents such as turpentine or suitable petroleum distillate. It seems probable that traces of metallic soaps formed by chemical interaction of pigment and oil not infrequently serve as protective emulsoids (p. 175), aiding in forming and stabilizing the pigment dispersion.

Dutch Process White Lead comes from the pots in which it is formed as relatively hard lumps. These are best ground in water which minimizes dusting and the consequent dangers from toxicity. The ground pigment can be settled to a relatively thick paste. If this aqueous suspension of pigment is agitated in a mixer with linseed oil, the pigment particles leave the water and enter the oil, the displacement being produced by surface tension effects. So complete is the exchange that the dispersion of pigment in oil contains only one or two per cent of moisture. Owing to the density of the pigment the suspension in oil is heavier than the water and settles to the bottom. The water withdrawn from above it is opalescent with pigment and is recycled to the grinding operation. A similar technique has been applied in recent years to colored pastes precipitated in aqueous solutions. These are "flushed" into oil or even into

varnish and thus used by the paint and enamel manufacturer. "Flushed" colors, while not as finely dispersed as well-ground pigments, eliminate the necessity of drying and regrinding in the vehicle, and are finding a place of increasing importance in the industry.

Calcimines or Distempers.

These are essentially water paints, water with some adhesive such as casein or glue forming the vehicle. Because the particles are not held in a medium of high refractive index, materials with a much lower refractive index, such as whiting, may be used as pigments with a corresponding reduction in cost. Cold-water paints are manufactured by grinding the pigments and casein or glue together in the dry state, a little boric acid, salicylic acid, or other preservative being added to prevent any mold formation. If casein be used as a binder, the mixture must be alkaline, the free alkali peptizing the casein. Such mixtures disperse readily in water and when applied to a surface produce a matte finish, which, after a few days' exposure to the air, becomes quite insoluble and adheres firmly to the underlying material.

BIBLIOGRAPHY

AM. SOC. FOR TESTING MATERIALS, *Symposium on Paint and Paint Materials*, March 6, 1935.

BARRY, T. H., DRUMMOND, A. A., and MORRELL, R. S., *Natural and Synthetic Resins*, D. Van Nostrand (1926).

BARRY, T. H., and DUNSTER, G. W., *Varnish-making*, Leonard Hill, Ltd., London (1934).

BOGUE, R. H., *The Chemistry and Technology of Gelatin and Glue*, McGraw-Hill (1922).

DURRANS, T. H., *Solvents*, D. Van Nostrand (1938).

EYNON, L., and LANE, J. H., *Starch—Its Chemistry, Technology and Uses*, W. Heffer and Sons, Ltd., Cambridge, England (1928).

FISCHER, E. K., *Ind. Eng. Chem.*, **33**, 1465 (1941).

GARDNER, H. A., *Paint, Varnishes, Lacquers, and Colors*, Institution of Paint and Varnish Research (1935).

GARDNER, H. A., *Physical and Chemical Examination of Paints, Varnishes, Lacquers and Colors*, Institution of Paint and Varnish Research (1939).

HEATON, N., *Outlines of Paint Technology*, Ed. 2, Lippincott (1940).

KRUMBHAAR, W., *Chemistry of Synthetic Surface Coatings*, Reinhold (1937).

MORRELL, R. S., *Varnishes and Their Components*, Henry Frowde, and Holder and Stoughton, Ltd., London (1923).

MORRELL, R. S., and DE WAELE, A., *Rubber, Resins, Paints and Varnishes*, D. Van Nostrand (1920).

OIL AND COLOUR CHEMISTS ASSOCIATION, *Varnish Making*, Chem. Pub. Co. (1940).

PARRY, E. J., *Gums and Resins*, Pitman & Sons, Ltd., London (1918).

PARRY, E. J., *Shellac*, Pitman & Sons, Ltd., London (1935).

RIDEAL, S., *Glue and Glue Testing*, Scott, Greenwood and Sons (1926).

NOTE. Important material is found in the files of the *Oil and Colour Chemists Association Journal*, London.

Chapter XV

Paper

Wood.

Wood consists predominantly of a mass of cellulose fibers, bonded together in a matrix of a cementing material called lignin. Wood from different species of trees differs both in the dimensions of the fibers and in the nature of the lignin around them. The fibers are always of an order of magnitude shorter than those used for textile purposes, the average length from the conifera (spruce, fir, etc.) being about 0.15 in. and from deciduous trees, such as poplar and maple, about 0.06 in. Most woods contain about 60% cellulose. The remarkable similarities between woods of a given general type, *e.g.*, the conifera, are brought out in Table I. The chief differences show up in the minor constituents, such as ether extract (largely terpenes) and the alcohol-benzene extracts (including the resins).*

In certain trees, especially spruce and balsam fir, the fiber bundles are sufficiently parallel and are held in such a way that small pieces of the wood of almost fiber dimensions can be split off by a purely mechanical process. For this purpose bolts of wood free from bark are pressed hard against a rotating wheel, usually of coarse carborundum, over which flows a stream of water. Small fragments approximating fibers are thus torn off and when screened from coarse particles form "ground wood," or "mechanical wood pulp," used in newsprint in high amounts, *e.g.*, 85% of the finished sheet, together with some chemical fiber to increase strength. Owing to its content of relatively unstable lignin, this pulp is not durable, and paper containing it deteriorates rapidly.

---

* Wood contains both pentosans and hexosans (*i.e.*, polyoses of the pentoses and hexoses). These vary in reactivity, both among themselves and relative to each other. Thus, the pentosans are preferentially dissolved, presumably by hydrolysis, in the bisulfite process, whereas they apparently resist solution in the alkaline processes, *e.g.*, Kraft. On the other hand, the more reactive hexosans are dissolved in the latter process, with the result that the weight yields of purified fiber are about the same in the two processes.

While the proportion of lignins differs somewhat with the species of tree, it increases somewhat with age.   The structure of lignin is uncertain, but it is known to contain both unsaturated

TABLE I.   VARIATIONS IN ANALYSES OF TYPICAL
CONIFEROUS PULPWOODS

	Low Values	High Values
Hot water extract	2.5	3.5
Ether extract	0.3	4.0
Alcohol-benzene extract	2.0	6.5
Hot 1% NaOH extract	12.0	15.0
Pentosans	6.0	14.0
Lignin	28.0	31.0
Cellulose	58.0	62.0

double bonds and aromatic nuclei with directly connected methoxy groups.   These groups appear to be parts of larger units, one of which may be the guaiacol-glycerol radical:

$$-O-\!\!\!\left\langle\bigcirc\right\rangle\!\!\!-CH(OH).CH(OH).CH_2-O-$$
$$CH_3-O$$
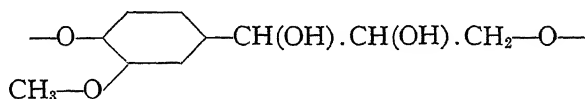
This may be linked with glucopyranose residues.   The presence of ketonic groups is less certain, although lignin apparently combines readily with sulfurous acid.   It is not attacked by acids other than sulfurous, but dissolves readily in hot dilute alkali with the formation of salts of soluble acids.   These properties are exploited in dissolving the lignin of wood from the cellulose fibers imbedded in it, taking advantage of the far greater inertness of cellulose.   The product is known as chemical wood pulp (see footnote, p. 342).

The three major methods of production of chemical pulp (Table II) are the sulfite process, employing a solution of sulfurous acid and acid sulfite; the soda process, using caustic soda; and the sulfate process, a modification of the soda method employing a mixed solution of sodium hydroxide and sulfide. Unfortunately, the cellulose is not entirely inert or unreactive to any of these reagents.   If the temperature be excessive, the concentration too high, or the time too long, cellulose fibers will be attacked and come out both shorter and weaker.

Whichever process is adopted, bark interferes and must be removed. This is best done by tumbling the wood, cut into lengths of about 4 ft., in a stream of water in large, rotating

TABLE II.   WOOD PULP PROCESSES

Name of Product	Method of Treatment of Wood	Types of Wood Used	Remarks
Mechanical wood pulp	Mechanical disintegration by grinding	Spruce and other coniferous woods	Used for newsprint; poor life
Sulfite pulp	Digestion with SO_2 soln. containing $Ca(HSO_3)_2$ and $Mg(HSO_3)_2$	Spruce, hemlock, fir, and other coniferous woods	
Soda pulp	Digestion with NaOH solution	Poplar, cottonwood, and other deciduous woods	Soft, glossy pulp
Sulfate pulp (Kraft)	Digestion with NaOH and Na_2S	Pine and other coniferous woods	Good for resinous woods; difficult to bleach; strong fiber

steel cages. The larger logs are split; all are then fed to chippers, in which the end of the log is pressed at an angle of 45° against knives bolted to a disk rotating at high speed. Uniformity in chip size, which averages 1/2 to 1 in. by 1/8 to 3/16 in., is attained by screening, finer material being discarded for fuel.

Sulfite Process.

From the point of view of tonnage and the general utility of the fiber produced, the sulfite process is the most important method of recovering cellulose from wood. It is adapted to coniferous woods of low resin content, e.g., spruce, fir, and hemlock.* The process consists in subjecting the chips to the action of a sulfurous acid solution which dissolves the lignin, probably by addition to the unsaturated linkages present to yield a sulfonic acid. However, because at the temperatures employed strong acids, such as sulfuric (sure to be formed by oxidation) and the sulfonic acids formed in the reaction, cannot fail to injure the fiber by hydrolysis, one must provide against their

* It is beginning to be employed for deciduous hardwoods, such as aspen and birch.

presence.   This is done by using some bisulfite (lime or mag-
nesia), which by its buffer action prevents excessive lowering
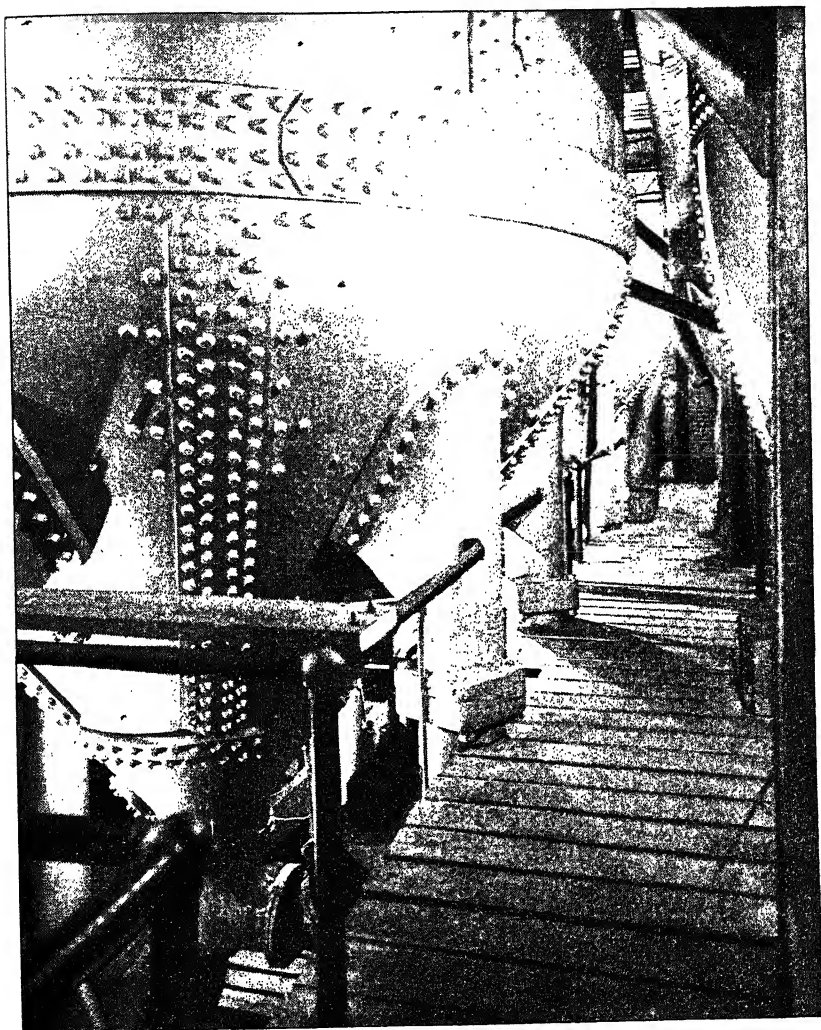
Fig. 1.   Bottom Cone of Digester Showing Blow Valve.*

of pH.   The metals presumably also serve to hold acidic de-
composition products in solution.   In practice, the solutions
used for digestion contain calcium and magnesium bisulfites in

* Courtesy of Brown Company, Berlin, N. H.

a quantity such that about 1% combined and 4 to 6% free sulfur dioxide are present.*   About 250 to 270 lbs. of SO₂ and 90 lbs. of calcium oxide are required per ton of wood, the excess SO₂ coming from recovery (see below).

Reaction rate is dependent upon temperature and the concentration of sulfurous acid.   The former is controlled by steam
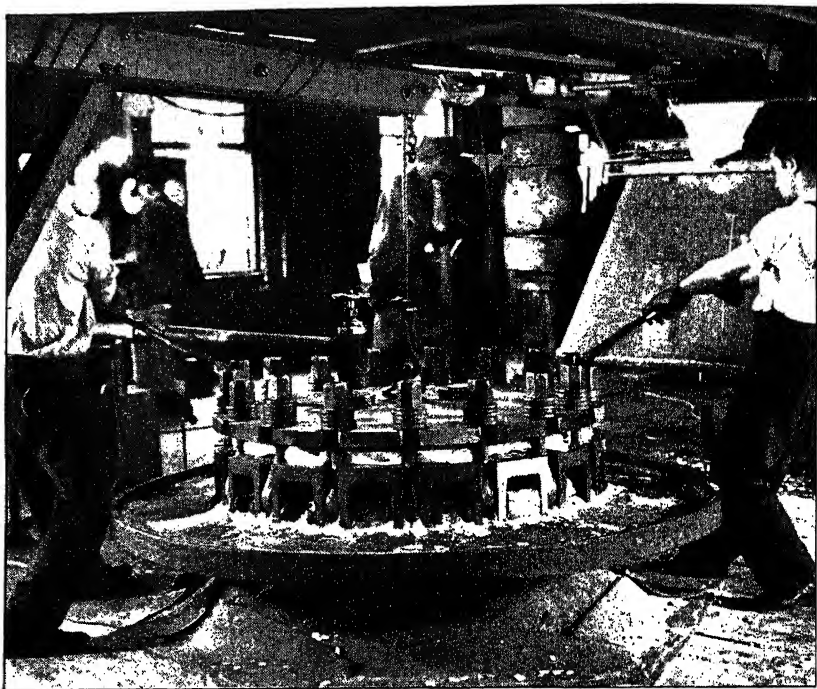
FIG. 2.   Bolting on Cover of Charged Digester.†

injection and the latter by pressure regulation corresponding to the temperature level, achieved by blowing off excess gas, largely SO₂, often accompanied by some liquor.   This relief makes it possible to get adequate control of both temperature and acid concentration as functions of time.‡

* Free SO₂ is that driven off by boiling; hence, combined SO₂ is that equivalent to neutral sulfite.   Consequently, a solution of 1% combined and 4% free sulfur dioxide is one containing 2% in the form of a mixture of calcium and magnesium acid sulfites, plus 3% as free sulfurous acid.

† Courtesy of Brown Company, Berlin, N. H.

‡ That control is not absolute is due to the presence of the bisulfites in somewhat variable concentrations.

The reaction vessel, called a digester, is a cylindrical shell about 15 ft. in diameter and 50 ft. high, made of heavy boiler-plate riveted or welded in sections, with a dome-shaped top and a cone-shaped bottom (Figs. 1 and 2). It is lined with

SULFUR OR PYRITES
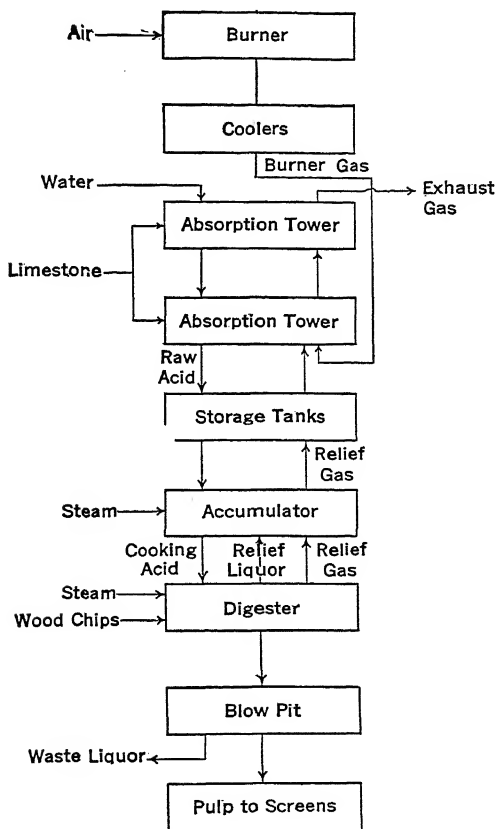
FIG. 3. Sulfite Process Flow Sheet.

acid-proof brick bonded with acid-proof cement, set in a layer of Portland cement. When the digester is filled with chips, the acid liquor is run in and heated, as by the direct addition of live steam. Operating temperatures are about 140° C. and pressures are of the order of 80 lbs. per sq. in. The concentrated sulfur

dioxide relief gas, together with any accompanying liquor, is introduced into a pressure accumulator (Fig. 3) which serves to preheat the fresh bisulfite liquor on its way to the digester, thus shortening the time of the cook itself. The gas from the accumulator is used as indicated to fortify the bisulfite liquor. The "cook" usually takes from eight to twelve hours, after which the digester contents are forced into a "blow-pit," and the pulp washed free from the dark brown, strongly sulfitic waste liquor. Many attempts have been made to utilize this liquor with some promise of economic success, in addition to avoiding river pollution. Since about 50% by weight of the wood and much of the sulfur used appear as soluble compounds in the liquor, the total quantity wasted is large.

The sulfite cook dissolves most of the pentosans. While the cellulose loss is small, there is normally some hydrolysis. Thus, a preliminary thorough impregnation of the chips with a 10 to 12% solution of sodium sulfite, followed by a cook in sulfurous acid of normal strength but for roughly half the time, results in a marked increase of the cuprammonium viscosity of the cellulose produced, indicating greater molecular chain length.*

Soda Process.

The soda process is advantageous for deciduous trees, such as poplar. The resulting fiber, short, soft, and easily bleached, has suffered little degradation but is high in pentosans, and some cellulose evidently dissolves. It is used, for example, in high-grade book paper with matte finish. The digestion of the chips is similar to that in the sulfite process, temperatures of approximately 170° C. and cooking times of four to six hours being employed, but the steel digesters require no lining. The treating agent, a 7 to 8% caustic soda solution, dissolves the lignin and other impurities, forming a "black liquor" containing the decomposed lignin as sodium salts of complex organic acids.

---

* Such a preliminary impregnation of the chips with buffer makes it possible to produce a high-grade pulp from hardwood, which by the conventional process yields a degraded pulp unless undercooked. It would appear that owing to the density of these chips the acid in an ordinary bisulfite liquid diffuses into them ahead of the buffer salts, resulting in degradation of the chips. This obviously cannot occur if the buffer is introduced ahead of the acid.
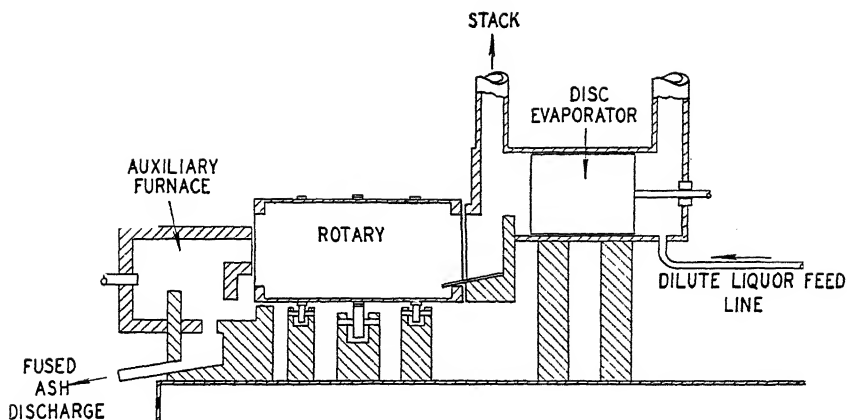
FIG. 4.   Black Ash Furnace.
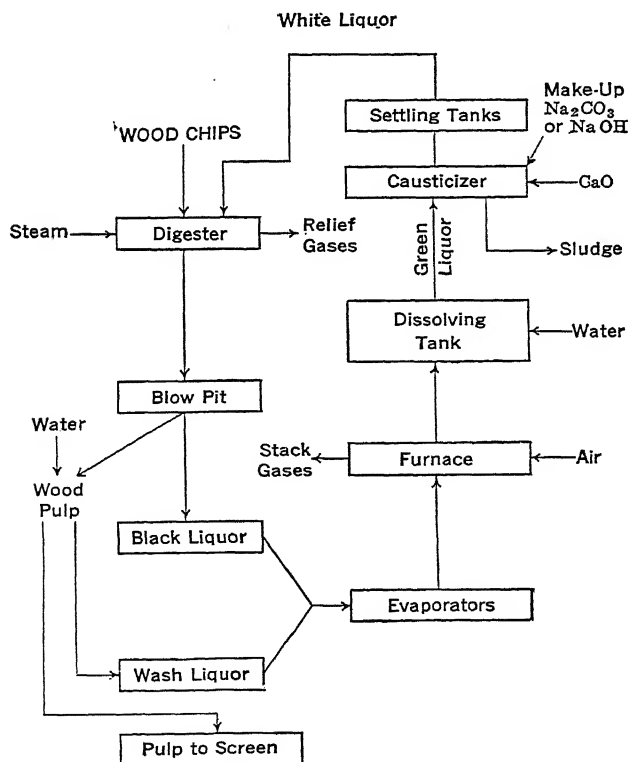
FIG. 5.   Soda Process Flow Sheet.

After reaction is complete the pulp is blown out, washed, and screened. The combined black and wash liquors are concentrated in multiple-effect evaporators to a thick syrup, containing about 50% solids, which is evaporated to dryness and finally incinerated in furnaces (Fig. 4). The oxidation of the organic matter in the liquor itself supplies a considerable portion of the heat. Furnaces in which the liquor is sprayed directly into the hot combustion gases are coming into use. The "black ash" from either type of furnace contains the sodium largely as carbonate, contaminated with unburnt carbon. The carbonate is dissolved, fresh soda is added to compensate for the not inconsiderable losses (10% and above), the solution is causticized with milk of lime and returned to the process, as indicated in Fig. 5.

Sulfate Process.

The sulfate process, a modification of the soda process, is applicable to all types of wood but is generally employed for the long-fibered conifera, being especially serviceable in treating highly resinous woods such as Southern pine. By suitable control a soft, long-fibered pulp of great strength, but retaining considerable brown coloring matter which resists bleaching, is obtained. From such fiber wrapping paper of superior strength, commonly known as Kraft, is made.

The treating agent present in major amount in the liquors of the Kraft process is caustic soda, but its action is tempered and modified by introducing along with it a considerable amount of sodium sulfide. This weaker alkali increases the total alkali stoichiometrically available in the solution without appreciably increasing its activity and the risk of injury to the fiber. The presence of the sulfide makes it possible to increase somewhat the temperature of cook, reduce its time, and still have a milder action on the cellulose. The effect of the sulfide is certainly not limited to buffer action, for its presence results in evolution of mercaptans and organic sulfides. Furthermore, its use makes it possible to get better solution of the non-cellulosic constituents of the wood than can be achieved by the soda process on the same wood. On the other hand, the cooking effects of sulfide

and caustic are sufficiently equivalent so that the sum of the two, each expressed in suitable equivalents, *e.g.*, as Na_2O, is spoken of as the "active alkali" in the cooking liquors. It is not advantageous to use too high a ratio of sulfide to caustic; this ratio (both constituents expressed as equivalents) is called the "sulfidity" and is usually kept in the neighborhood of from 25 to 30%. The sulfide is in practice accompanied by a nearly equal amount of sodium carbonate, which, however, seems to be relatively inert.

The Kraft cycle of operations is similar to that of the soda process (Fig. 5). The sulfide is not introduced as such but as sodium sulfate. The black liquor which is recovered from the digester and from washing the pulp, as in the corresponding operations of the soda process, is evaporated and the make-up sodium sulfate necessary to replace the losses is usually added prior to the initial stages of incineration.* The black liquors are carried to dryness and the organic matter decomposed in a furnace operation substantially identical with that of the soda process. However, the formation of sodium sulfide requires reduction of sodium sulfate by the carbon of the organic matter, a reaction that occurs only at temperature levels so high that it is necessary to smelt the black ash from the Kraft operation, a step which is not essential in the soda process. In consequence, the auxiliary furnace of Fig. 4 can never be omitted in sulfate production and must be operated at an adequately high temperature level. Because of the high heat consumption of reducing sulfate to sulfide with carbon (approximately 45 kg. cals. per gm. mol even when the carbon oxidizes completely to CO_2) it is necessary to have an adequate fuel supply at this point.

The smelt is dissolved as "green liquor" and causticized to "white liquor" as in the soda process. Because losses of sulfur in connection with both digestion and furnacing exceed those of soda, the atomic ratio of sodium to sulfur in the treating liquor is approximately 6 or 8 to 1, compared with the value of 2 to 1 in the make-up sodium sulfate.

---

* It has become increasingly difficult in recent years to get adequate supplies of sodium sulfate, so that other sources of replenishment of the losses in the cycle have come into use. For example, a mixture of sulfur and soda ash has been employed.

Bleaching.

Pulp is freed of incompletely digested chips or knots by washing through screens and flowing through riffles (long troughs containing baffles) in which the coarser impurities settle. This is best done after considerable dilution of the pulp. The product is always dark-colored and for white paper must be bleached.* This may be accomplished by the use of a hypochlorite solution which, in the so-called one-stage bleach, is usually the calcium salt. The pH of such a calcium hypochlorite solution, initially 11 to 12, falls steadily in the presence of pulp but should not be allowed to go below 8.

The coloring matters left in the pulp act like lignin in that they will add chlorine at low pH and are thereby converted into alkali-soluble derivatives. Furthermore, if the solution is sufficiently acid, degradation of the fibers due to oxy-cellulose formation (p. 154) becomes very small. While it is true that such a treatment, even where the chlorination is followed by an alkaline wash, reduces the color relatively little, a final treatment with a dilute calcium hypochlorite solution at a pH above 8 can clean up the color very effectively. These facts are the basis of the so-called two-stage bleach. The initial treatment is carried out with chlorine, which, because it tends to react with water to form hypochlorous and hydrochloric acids, gives a pH which may be as low as 2. The alkali solubility imparted to the impurities by the chlorine absorbed by them in this initial treatment can result in a considerable saving in the total chlorine consumption required for bleaching (20 to 40%). This technique makes it possible to treat even Kraft fibers hitherto considered commercially unbleachable.† Where the ligneous material is large in

---

* In a certain sense the bleaching may be looked upon as a continuation of the cooking process. Lignin can be removed from the original wood by chlorination, followed by solution of the chlorinated products in alkali sulfite, but the chlorine consumption is excessive. The various methods of cooking represent alternative ways of removing the lignin and other impurities. It is possible to remove much of the lignin during cooking with very little injury to the fiber; the degradation increases considerably, however, as one attempts to approach complete purification by cooking alone. Purification can be completed by proper bleaching technique with little or no injury to the fiber, although at greater expense per unit of lignin removed. The distribution of the purification between cook and bleach is a matter of economic balance depending on the quality required for the specific use to which the product is to be put.

† The cook may also be modified to some degree, e.g., its time increased.

amount, it may be desirable to use the treatment with chlorine followed by dilute alkali repeatedly prior to the final bleaching.

It is imperative to eliminate residual hypochlorite. This is sometimes done by the use of "anti-chlors," such as sulfites or sodium thiosulfate. On the other hand, it is usually preferable to rely on careful washing. If the pH of the operation be properly controlled, time and temperature of treatment kept within reasonable limits, and the reactants adequately removed, degradation of the fiber in bleaching is negligible. In consequence, mechanical strength of the paper made from it will be unimpaired.

Structure of Paper.

Cellulose fibers, separated and purified as just described, are amorphous in structure, permeable to most liquids, highly resistant to chemical and physical action, and of extremely low plasticity, *i.e.*, separate fibers do not readily adhere or coalesce. They can be dissolved by various reagents (p. 366), all of which, however, apparently involve the cellulose in relatively deep-seated chemical transformation. If the fibers be submitted to the incipient action of any of these reagents they are swollen and plasticized, and coalesce readily when brought into contact.

Fig. 6. Unbeaten Linen Fibers (70×).

The fibers are remarkably inert to the action of non-polar liquids, but respond to the action of highly polar ones, swelling in them to a degree determined mainly by their polarity, as brought out

in the first two columns of Table III.   Perhaps the most impor-
tant characteristic is the fact that if the swollen fibers be beaten,
i.e., macerated and mauled mechanically, as, for example, in a
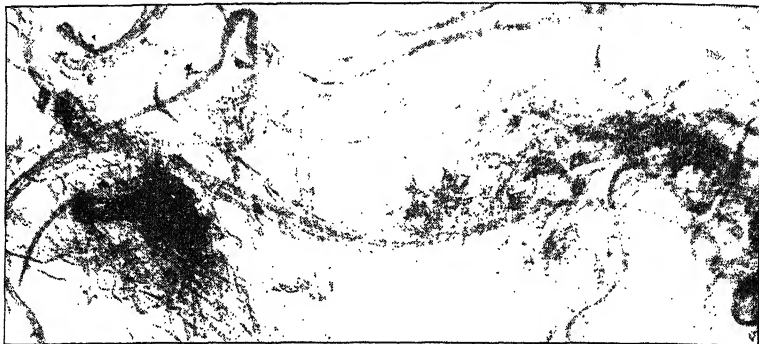
FIG. 7.   Beaten Linen Fibers (70 ×).

ball mill while suspended in a polar liquid, the softened walls
rupture and hundreds of little fibrils are set free by a frazzling
action (see Figs. 6–7).   The fibrillation is evidently made
possible by the mechanical weakening of the fiber by the swelling
it undergoes in the solvent, since the extent of the fibrillation,
in contradistinction to a mere cutting or shortening of the original
fiber, is definitely related to the swelling effect of the solvent on
the original, unmacerated cellulose, and no substantial degree of
fibrillation can be developed in non-swelling, inert liquids.

A dilute suspension of cellulose fibers in water can be strained
through a cloth or screen, the fibers being filtered out as a matte
or sheet which, when dried, forms paper.   By imparting a sharp
vibratory motion, in a direction parallel to the surface of the
screen, as the water is flowing through it, the fibers are more
effectively interlocked and the sheet correspondingly strength-
ened.   In the early centuries of paper manufacture the sheet
was always formed by hand in this way, but present-day
technique is mechanical.

If a mass of cellulose fibers, from any of the sources mentioned
above, be formed into a sheet without the mauling and macera-
tion already described, a porous, open matte is obtained which is
similar in structure and characteristics to blotting paper.   If,

however, the fibers, while suspended in water, be macerated by mild mechanical action for a considerable period of time, not only does one observe the fibrillation illustrated in Fig. 7, but the final sheet of dry paper formed from the mass is far stronger, denser, and, instead of being opaque, can be translucent to the point of transparency. The remarkable transformation in properties induced by mechanical mauling (beating) under water lies at the heart of the paper-making process.

Adequately to understand the transformation requires the study of further facts. In the first place, one must keep in mind that usually the opacity of a sheet consisting of small particles increases as particle size decreases.* In complete contrast to this, the fibrillation accompanying beating, instead of increasing opacity of the final sheet, very greatly decreases it. Apparently beating, despite the reduction in particle size, causes an elimination of particle surface in the ultimate sheet, and this in turn can only result from coalescence of the particles on drying. This explanation is confirmed by the greatly increased resistance of the sheet to flow of water or other fluids through it, advantage of which is taken in an important method for measuring the transformation achieved by beating, namely, the determination of freeness.†

Coalescence of the fibers of a beaten stock on drying is confirmed not only by the increase in the mechanical strength of such a sheet with time of beating (Fig. 8), but also by its behavior toward various liquids. Thus, the mechanical strength of a sheet of paper is always greatly reduced by soaking it in water. This reduction is not primarily due to a mere mechanical lubrication of the fibers by the liquid so that they slip past each other, for the influence of non-aqueous liquids on the strength of a sheet is far different from that of water. This is brought out in Fig. 9, which shows the tensile strength of samples of waterleaf paper ‡ differing widely in degree of beating, as influ-

* Provided one does not get down to dimensions approximating the wave length of light (see p. 133).
† The fiber is suspended in water at a definite concentration and the water allowed to drain through a standard screen under definite conditions of head. The rate of drainage, expressed in arbitrary units, is the freeness of the stock. Two stocks substantially identical in fiber length may differ many fold in freeness.
‡ Essentially pure cellulose.

enced by mixtures of alcohol and water of different concentrations. It will be noted that water always weakens a sheet much more than alcohol. Beating increases the strength of the sheet,
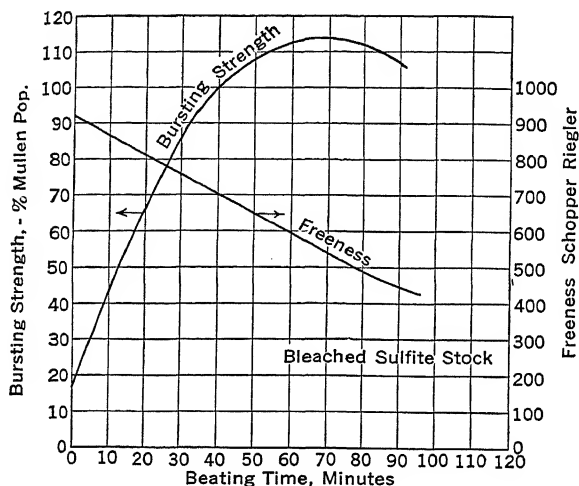
FIG. 8. Influence of Beating Time under Constant Conditions of Roll Setting, Stock and Temperature.

(The Mullen tester stretches the sheet across a hole in a test plate, with a thin rubber diaphragm beneath the paper. Air pressure is built up beneath the sheet and the pressure necessary to "pop" is determined. In this diagram strength at 40 minutes beating time is taken as standard).

but while the absolute difference in tensile strength resulting from the action of alcohol on the one hand and water on the other is greater in the highly beaten stock than in the relatively unbeaten material, the ratio of the two effects is not greatly different. On the other hand, while even absolute alcohol weakens the sheet considerably, an inert non-polar liquid like benzene weakens it little, if at all, and under special conditions can actually increase its strength.

The influence of the liquid on the characteristics of the sheet parallels closely its swelling action on cellulose fibers (see Table III). While the effect of benzene is not given in that table, it undoubtedly differs little from that of the heavier hydrocarbons comprising fuel oil.

The characteristics of cellulose fibers can be even more clearly appreciated by the action of liquids upon them in other forms.

TABLE III. BEATING OF A PAPER STOCK IN VARIOUS LIQUIDS*

Reagent	Solvent Polarity, Expressed as Dielectric Constant	Comparative Swelling (% increase in volume of fiber)	Freeness of Beaten Pulp [a]	Character of Sheet
Formamide	84	125	410 after 2 hrs.	Strong [b]
Water	80	90	90 after 4 hrs.	Strong
Ethylene glycol	41.2	89	140 after 1.5 hrs.	Strong
Methyl alcohol	34	62	340 after 2 hrs.	Fairly strong
Butyl alcohol	18	4	800 after 3 hrs.	Weak
Amyl chloride	6.5	4	800 after 4 hrs.	Very weak
Fuel oil	2.0	2	825 after 8 hrs.	Very weak

[a] Data at equal times of beating are unfortunately not available.
[b] Mechanical strength also accompanied by oil and water resistance.

Thus, if an ordinary thread or cord of cotton fibers be soaked in water, its tensile strength is decidedly increased, whereas if the same cord be soaked in absolute alcohol, or, under certain con-

Curve	Time of Beating, Hours	Dry Strength, Lbs. In. of Width
A	5.0	27.5
B	4.5	24.7
C	3.0	18.8
D	1.5	3.5
E	Unbeaten	1.5

Mol. % Ethyl Alcohol in Conditioning Solution-Other Component Water
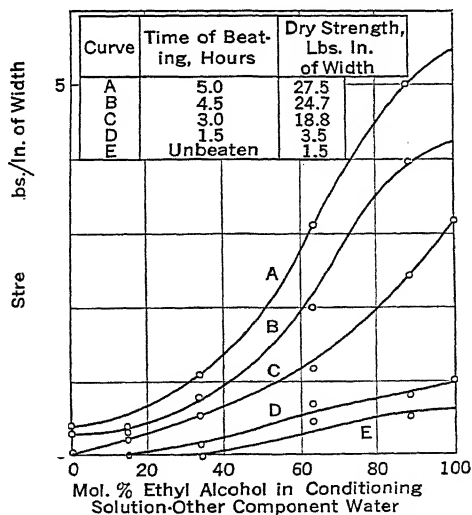
FIG. 9. Tensile Strength of Paper in Water-Alcohol Mixtures.†

ditions, in benzene, tensile strength is reduced. It will be noted that in the case of water the change in strength is exactly

* KRESS, O., and BIALKOWSKY, H., *Paper Trade Jour.* 93, No. 20, 35 (1931).
† KIRKPATRICK, W. A., *Mass. Inst. Tech. Thesis* (1932).

opposite to that which would have been observed had this same cotton fiber been converted to paper stock by beating, and formed into a sheet of paper by matting and drying. The same may also be true in the case of benzene.* Additional light is thrown on the situation by the fact that the reduction in strength of thread or cord resulting from the action of absolute alcohol or benzene is practically identical with that produced by drying the thread to a bone-dry condition.

The key to these apparently anomalous phenomena is the fact that, whereas cellulose is plasticized by water, the plasticizing action of alcohol is much less and that of benzene is negligible. The water renders the surface of the cellulose fibers soft and slippery; alcohol does this in less degree; on the other hand, the surface of the dry fiber is harsh and the non-polar benzene exerts no softening or plasticizing action upon it. The strength of the cotton thread, while due to the strength of the ultimate fibers, is modified by the fact that the friction between them, especially when drawn taut, is sufficient to prevent much slippage of one fiber past another. In the case of the air-dried cord, there is a little slippage of the fibers as the load is applied, but not enough to equalize the load on the individual fibers, some of which reach their breaking point before the others are fully loaded. Consequently, the cord breaks piecemeal, the rupture of the first fibers throwing the load on the others. When the cord is soaked in water the surface of the fibers is plasticized and rendered slippery, so that as the load increases the slippage adjusts the strain of the individual fibers. However, the fibers do not pull out; first, because their length is sufficient to give a large surface of contact and the twist produces adequate friction through sidewise pressure of the fibers upon each other, and, second, because the plasticization of untreated cellulose fibers by water is relatively slight. The additional adjustment of the load among the various fibers in the cross section of the thread caused by the incipient slippage is enough to increase the strength of the thread in some cases as much as 40%. If the fibers are impregnated with benzene and exposed to the air before or during testing, the moisture distilled out by the evapo-

* In all cases the effect of benzene is small relative to that of water.

ration results in dehydration of the fibers, hardening them, and, despite any lubricating action of the benzene, lessens the slip and therefore reduces the equalization of load between the fibers and lowers the strength. That the strength reduction is due primarily to elimination of water from the fibers is indicated by the fact that substantially the same reduction results from ordinary drying. It is not entirely clear why soaking in absolute alcohol reduces the strength of cord to almost exactly the same extent as drying or soaking in benzene; it may be that the alcohol, unlike water, has little or no plasticizing action on the outer surface of the natural fiber, although it can plasticize the surface of the fibrils, due to the presence on them of polar linkages released by the disruption of the original fibers by beating, as seen from Fig. 9.

The conclusion seems inescapable that the surface of the fibrils produced by the subdivision of the initial cellulose fibers by beating differs from the surface of the original fiber in being readily plasticized by water. Unlike the original fibers, the wet surfaces of the fibrils, brought together under pressure and dried in this condition, coalesce to a considerable degree, forming a relatively strong bond. This plasticization of a fiber by beating is called *hydration* by the paper maker.* It is evidently the major cause of both the decrease in freeness of the stock on beating and the increase in strength of the paper sheet formed from it, as indicated in Fig. 8.

The reason for the low plasticity of the original cellulose fiber as compared with that of the fibrils is not clear. However, it may be that in the synthesis of the fiber in nature by the laying down side by side of a multitude of molecular chains of cellulose, the highly polar hydroxyl groups along the chain exert a sufficient mutual attraction to cause a rotation of the molecules, which would result in a more or less complete neutralization of the fields of force around the groups. Such an orientation

---

* In the bisulfite process a mild cook, involving low temperature, time and concentration, gives a fiber which hydrates more readily on beating and is used to form dense and semi-transparent papers. While this may be due to the high percentage of pentosans in the fiber, it is more probably due to the avoidance of fiber densification which accompanies higher temperature of cook in more concentrated solutions (see p. 175). One can take a given original pulp and by different treatments (different methods of bleaching, alkali cooks, and the like) produce products identical by conventional tests, but differing widely in beatability (capacity for hydration in beating) and chemical reactivity.

would tend to reduce to a minimum any outwardly directed hydroxyl groups on molecules located on the surface of the fiber, with a corresponding reduction in the activity and consequently the plasticity of that surface. Fibrillation in beating inevitably involves the exposure of molecular surfaces which in the original fiber were in immediate contact with other cellulose molecules. Such surfaces would thus be expected to have a relatively high susceptibility to reaction with water. Chemical conversion of cellulose to soluble derivatives indicates that beating has little or no effect on the structure of the ultimate cellulose molecule itself, *i.e.*, chain length is probably not affected.

Beating.

As a source of cellulose, the old paper makers used cotton and linen rags * for which the mechanical maceration process referred to above was very necessary. For this purpose, the Dutch paper makers devised a machine which, somewhat modified, is still used, namely, the Hollander or beater. It consists of an elliptical tub (Fig. 10) provided with a central partition extending along its length to within a few feet of each end. This forms a continuous channel, and in this channel the cellulose pulp suspended at about 5% concentration in water is made to travel by a rotating cylinder or "roll" placed in its path on one side of the mid-partition. In the periphery of this cylinder are

FIG. 10.   Beater or Hollander.

embedded a number of blunt bars or knives parallel to its axis. As the cylinder rotates these knives all but touch a corresponding series of knives set in a bed plate underneath the cylinder.

* Rags are sorted for quality and color; after removing buttons, etc., they are mechanically cut and dusted. They are then boiled for several hours under steam pressures of 15 to 60 lbs. in alkali, *e.g.*, 5–10% milk of lime, to which soda ash is often added, to loosen color, size, etc. The pulp is dumped, drained, washed thoroughly in a breaker (a beater somewhat modified to disintegrate the rags but minimize maceration of the fibers) with large quantities of water, and finally bleached in the same unit. It is then transferred to the beater, in and beyond which the operations parallel those with wood pulp.

The clearance is subject to exact regulation, and as the cellulose fibers are driven between the two sets of knives by the rotation of the cylinder, they are bruised, torn, frayed at the ends, and highly macerated. The length of the fibers may or may not be considerably changed by the operation, depending on the sharpness of the knives * and the clearance. Thus, by regulating the speed and clearance of the roll and the time of beating it is possible merely to cut the fibers, to fray and fibrillate them, or to pound them to a pulp according to the type of final paper which it is desired to produce. The specific cases cited might correspond, for example, to blotting, cigarette, and grease-proof papers, respectively.

It is well recognized that different pulps vary greatly in ease of beating. Thus, bleached sulfite pulps are easily hydrated in from one to three hours, while soda pulps are very free and resist hydration. Kraft or sulfate pulps are also easily hydrated. Hence, for glassine papers, pulps capable of high hydration are sought, while for soft, opaque, printing papers soda pulps are used, hydration being kept at a minimum and fraying and cutting of the fiber encouraged by suitable adjustment of the beater blades.† Hydration proceeds more rapidly at low temperatures and is favored by high pH.‡ In order to complete the beating process and insure that all fibers shall be as nearly uniform in length as possible, the beaten pulp is often passed through a "Jordan engine" or "refiner." This is a conical shell containing within it a concentric rotor, each fitted with somewhat blunt knives so arranged that when the rotor is in motion a sidewise shear is produced. The clearance between the two is carefully adjusted, and the long fibers are cut and brushed out as the pulp passes through. The tendency at the present time is to cut down beating time as far as possible, relying upon the refiner to complete the process.

* Blunt blades tend to macerate rather than to cut and shorten the fibers.
† There is some evidence that "beatability" is connected with the so-called hemicellulose content of the pulp, *i.e.*, quasi-cellulosic material less stable than cellulose itself. Addition of cellulose degradation products to a pulp increases its rate of beating.
‡ As will later appear (p. 360), caustic soda reacts with cellulose, as shown by the fact that it is picked up selectively from aqueous solution and, if sufficiently concentrated, will dissolve less resistant forms of cellulose. The effect of alkali on beating, therefore, may well be due to increased plasticization of the fiber (p. 280).

Sizing.

For many purposes, a sheet of paper produced from beaten fibers without addition agents suffers from certain defects. Thus, unless heavily beaten, it is very absorbent and has many loose fibers. These defects may be overcome by sizing, which is essentially the introduction of an amorphous bonding material between and around the cellulose fibers to cement them together and fill up the voids. Besides fulfilling this function, the size can develop water repellency and serve to bond the fillers, when added, to the paper. The presence of fillers also plays an important part in retarding the softening action of water.

Sizing may be introduced by impregnating the paper sheet (tub sizing) or, in engine sizing, by addition of the necessary constituents to the beater, which forms an ideal mixer. Water-soluble materials must be applied by the former method, since if introduced into the beater or into the stock before the sheet is formed on the paper-making machine they would be almost completely washed out and lost. In spite of the greater expense of tub sizing it is used for high-grade papers since an exceedingly smooth surface can be obtained. As a tub size, glue is almost ideal, because while readily wet with water it does not dissolve at low temperature so that its effective water repellency is considerable. Casein is even better because it can be applied in solution in such a way as to become insoluble or practically so on drying, as by evaporation of the ammonia used to dissolve it. Water glass has its value because the silicic acid not only serves as a bond but is dehydrated and becomes relatively insoluble on drying. However, sizing as a separate operation can be eliminated if an insoluble sizing material can be completely dispersed in the stock itself, so that as the sheet of paper is formed the particles of size are entrapped, retained, and function satisfactorily in the finished product. A highly satisfactory material of this type has been found in the so-called rosin size.

If rosin soap (p. 264) is allowed to react with aluminum sulfate in concentrated solution a curdy precipitate of the aluminum salt, of somewhat indefinite composition, is obtained. Satisfactory dispersion of this is impracticable. If, however, it be formed in very dilute solution, its dispersion is excellent

and it can be uniformly and satisfactorily distributed on the fibers. The resinate is usually precipitated in the beater itself by addition of the necessary amount of rosin, either dissolved as sodium salt or emulsified in water, followed by the alum or aluminum sulfate. The order of addition is sometimes reversed. The alum used is several fold that stoichiometrically equivalent to the rosin but a large fraction of it passes through the sheet and out in the waste water. The precipitate can be synthesized before adding to the paper stock and this is sometimes done. To obtain satisfactory sizing the pH in the beater should be in the neighborhood of 4.0 to 4.5; after dilution, as the stock flows to the paper-making machine, the pH rises to between 5.0 and 6.5.

Acid used alone instead of aluminum sulfate does not give a satisfactory size and of the metals only aluminum is completely satisfactory. Aluminum resinate evidently possesses the requisite bonding power and water repellency in a unique degree.*

While the mechanism of beater sizing is in dispute, it would appear to be, in no small degree, mechanical. This is indicated by the fact that the specific method of incorporation of the size has but a minor effect on the results. It is also confirmed by the phenomena of retention of the particles of loading material or filler, such as China clay and calcium sulfate, frequently added in the beater. Thus, whereas an unbeaten sheet retains but little of a given filler (Table IV), the introduction of alum

TABLE IV.   INFLUENCE OF PARTICLE SIZE OF CLAY FILLER
ON ITS PERCENTAGE RETENTION †

Character of Pulp	% Alum	5 μ	5–12 μ	12–20 μ	20–44 μ
Unbeaten and unsized	0	8	15	25	32
	2	36	34	47	54
Unbeaten; sized with 2.3% rosin	0.84	28	32	39	42
	2.5	38	41	49	55
Beaten and sized with 2.3% rosin	2.5	71	72	74	76

increases the retention. Addition of rosin still further runs up retention, as does beating of the stock. Decreasing the particle size of the filler reduces retention but, although the effect of

* Ferric iron is also good in this respect but is ruled out on the basis of color and oxidizing tendency.
† ALBERT, C. G., *Tech. Assoc. Papers* **20**, 286 (1937).

this factor is very marked on unsized stock, with a sized beaten stock it almost disappears.

Paper Machines.

The process of making paper from the prepared pulp by modern machines does not differ in principle from that employed in handmade paper, although the rate of production has increased until now a sheet of newsprint 320 ins. wide and 1400 ft. long can be made each minute. This has become possible by the perfection of a machine known as the Fourdrinier, a marvel of perfect co-ordination (Figs. 11 and 12). A suspension of
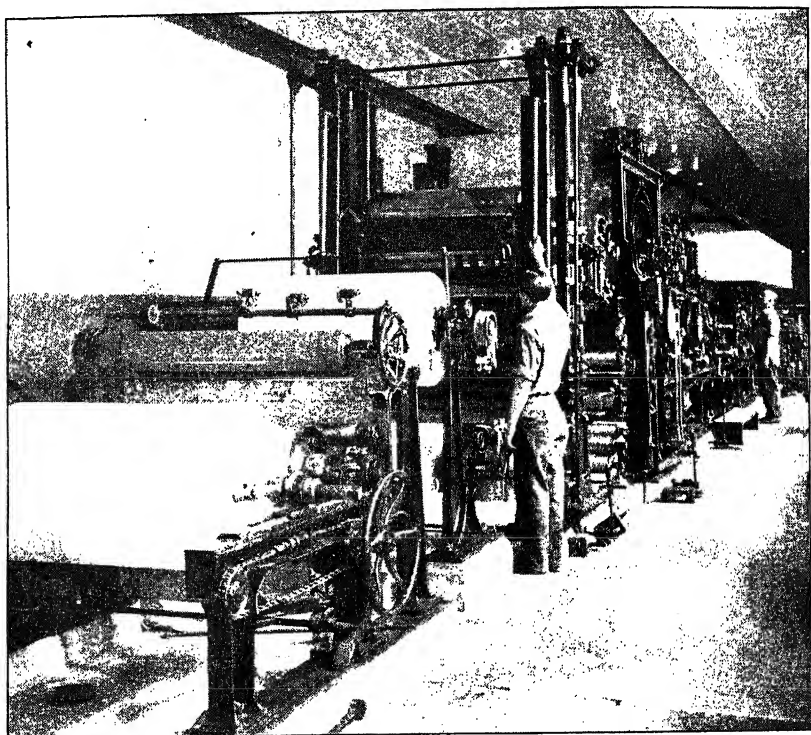
Fig. 11.   Experimental Paper Machine.*

fibers, filler, size, and such color as is desired is prepared as already described and diluted to a concentration of about 0.5%.

* Courtesy of Brown Company, Berlin, N. H.

This suspension is fed at a uniform rate upon a moving table of fine wire cloth, the first several feet of which is usually provided with side edges of travelling rubber belts called deckle straps. Usually the head end of the wire table is given a slight sidewise shake which helps to interlock and intertwine the fibers into the resultant sheet. Modern high-speed Fourdriniers use neither deckle straps nor the sidewise shake. The wire and its suspension of fibers pass first over a series of table rolls, where
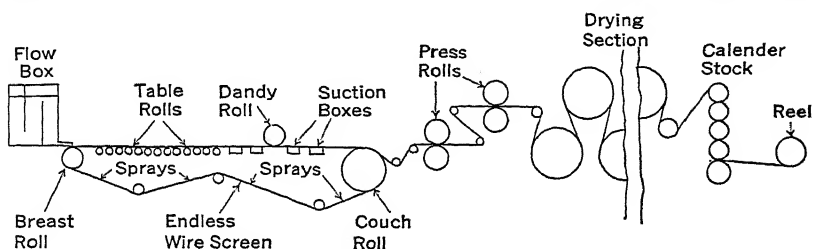
FIG. 12.   Paper-making Machine.

over half of the water is removed. They then pass over rectangular suction boxes, in which a vacuum is maintained and more water is sucked out. The wire then passes around a perforated roll called the couch roll, in which also a suction is maintained, and returns as an endless belt to the head end of the machine. The paper web is separated from the wire on the couch roll, and is taken up on another endless belt, this time of woolen felt. Further water is removed by press rolls, and the sheet of paper, now of 30 to 35% consistency * and possessing considerable mechanical strength, is transferred to a series of steam-heated cylinders about 4 ft. in diameter, and from 25 to 30 in number, on which it is dried. Not only must the different sections of the machine be perfectly synchronized, but the speed of the drying cylinders must be such as to allow for the shrinkage of the sheet as the moisture leaves it. A series of polished steel rolls give it a hard, dense surface.† If a highly polished

---

* The term *consistency* as employed in the industry is the per cent by weight of air-dry fiber (which in turn is defined as fiber containing 10% moisture) in a given suspension of fiber in water.

† Recently two-wire machines have been developed with the object of obtaining evensidedness and removing the fluffiness of the sheet on the wire side, objectionable in high speed, offset, lithographic printing. Two similar wires traveling in opposite directions are used, so that the sheets from them meet just before a press roll and their wire sides are consolidated together.

surface be desired the sheet is calendered by passing between rolls, one of which revolves at a higher rate of speed than the other.

Coated Paper.

For the perfect reproduction of very fine lines, such as half-tone prints, the fibers are entirely covered by a surface coating of clay or other pigments, held in a solution of glue or casein. The coating liquid is applied by dipping, excess removed by brushes or air blast, and the paper dried by passing over cylinders, in long festoons, or carrying in a horizontal position on an upward blast of air. The coating is burnished to a smooth surface by highly polished rotating rolls.

An important type of product, though of limited use, is made by plasticization of the ultimate fibers of a sheet of paper by treatment with a cellulose solvent under conditions avoiding solution but securing a high degree of coalescence. The commercially important reagent is $ZnCl_2$, used in about 75% solution at 40° C. Thus, pure cellulose sheet is drawn through the bath and rolled up on drums or mandrels, where it coalesces to an almost homogeneous mass. It is then carefully washed, preferably first by gravity displacement, dried, pressed, and calendered. Mechanically strong and tough, the product, so-called vulcanized fiber, is a valuable material for many severe service conditions, e.g., insulation of rail joints in automatic signal systems, gears, washers, etc.

BIBLIOGRAPHY

CLAPPERTON, R. H., and HENDERSON, W., *Modern Paper-making*, Blackwell (1940).

CROSS, C. F., and BEVAN, E. J., *Papermaking*, Spon and Chamberlain (1920).

Manufacture of Pulp and Paper, Vols. III, IV, and V, McGraw-Hill (1938).

RUE, J. D., *Paper Trade J.*, 114, T.A.P.P.I., 32–34, 1942.

SUTERMEISTER, E., *Chemistry of Pulp and Paper Making*, Wiley (1941).

WEST, C. J., *Bibliography of Pulp and Paper Making* (T.A.P.P.I.), Lockwood Trade Journal Co. Inc., New York (1938 and 1940).

WEST, C. J., *Classification and Definitions of Paper* (T.A.P.P.I.), Lockwood Trade Journal Co. Inc., New York (1928).

WITHAM, G. S., *Modern Pulp and Paper Making*, Ed. 2, Reinhold (1942).

Chapter XVI

The Plastic Fibers

Rayon.

The technique of manipulating an amorphous solid by precipitation from solution in a suitable shape is of particular importance in the processing of cellulose. However, the limitations imposed are even more severe than for the method of solution followed by evaporation, used for the production of plastic films. As in that technique, the material can be manipulated only in the form of thin films or filaments, because it is otherwise extremely difficult to obtain uniform precipitation and coagulation. Cellulose, by virtue of its three hydroxyl groups (p. 151), is capable of giving soluble esters or ethers, which can be precipitated in the desired shape, either as such or as regenerated cellulose.

The strength, beauty, and high cost of natural silk made attempts to develop a less expensive fiber resembling it in appearance and properties inevitable. In the last forty years it has become possible to obtain from cellulose an almost ideal material, resulting in the development of a large industry, the manufacture of artificial silk or rayon.

Because natural silk is a protein fiber of smooth surface with a uniform and approximately circular cross section (Fig. 10, p. 503), it has a highly prized luster or sheen. The first problem of substitution consists, therefore, in producing a smooth, uniform, cylindrical fiber of the proper size. The earliest attempt to make a cellulose thread simulating the appearance of silk which achieved any appreciable success, was the process of mercerization (p. 488). This produces a thread in which the fibers are cylindrical and much of the helical twist has been removed, and consequently develops a pronounced luster. However, the fibers are not entirely transparent and, in the ultimate thread,

some twist remains, both factors reducing luster.   Furthermore, there is no way of controlling the size of the fiber itself.   It proved impossible to produce satisfactory substitutes for silk until the development of methods for the complete mechanical synthesis of the fiber by imitating as closely as possible the actual technique of the silkworm in producing a natural silk.   The first step is the preparation of a solution of an amorphous solid suitable for use as a fiber.   Until recently, in all commercial processes cellulose or one of its derivatives was employed.   The cellulose solution is forced through an extremely small orifice to give a filament, and the plasticity of the cellulose is then destroyed either by its precipitation or by evaporation of the solvent.   A number of filaments produced simultaneously are
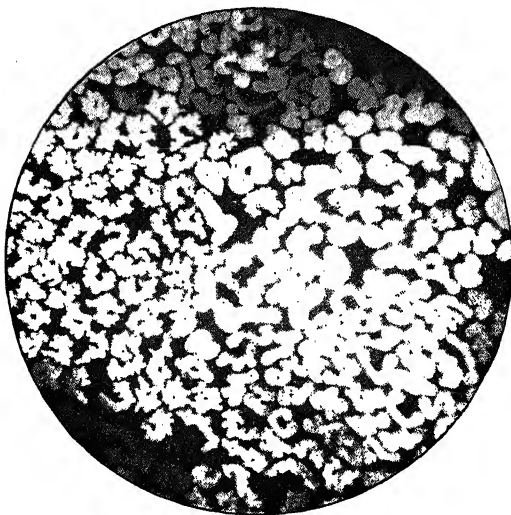
FIG. 1.   Usual Type of Viscose and Acetate Rayon Cross Sections.*

twisted together to form a thread, typical structures of which are shown in Fig. 1.   The threads must be subjected to various finishing operations.

The processes are conveniently classified in terms of the four derivatives employed to bring the cellulose into solution: (a) cellulose xanthate dissolved in water, the viscose process; (b)

* Courtesy of Prof. E. R. Schwarz, Textile Laboratory, Massachusetts Institute of Technology.

the cuprammonium complex, cellulose dissolved in aqueous solutions of ammoniacal copper salts; (c) nitrocellulose and (d) cellulose acetate, both dissolved in organic solvents. In all cases, the cellulose or its derivatives can be regenerated from the solution by suitable liquid precipitating agents, but in the last two the solvent is often removed by evaporation. While, from the point of view of consistency, this latter technique should be treated in an earlier section, it is covered here because of the complete parallelism of all processes in the earlier stages of manufacture.

About 75% of the rayon produced in the United States is made by the viscose process and almost all of the remainder is acetate, the importance of which is increasing rapidly.

Viscose Process.

The viscose process depends upon the fact that carbon disulfide, in the presence of caustic soda, will form with alcohols water-soluble, highly ionized salts, known as xanthates,

$$Na.OH + R.OH + CS_2 \rightarrow C\underset{S.Na}{\overset{OR}{\diagdown}}\!\!=\!\!S + H_2O$$
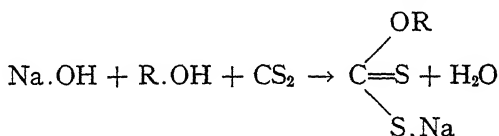
where R is the organic radical of the alcohol. With the poly-hydric alcohol cellulose the above reaction takes place easily, forming the sodium salt of cellulose xanthogenate or cellulose xanthate. It is water-soluble as an emulsoidal solution, the high viscosity of which gave the product its name. The solution is unstable, slowly but progressively hydrolyzing with ultimate regeneration of the cellulose. This hydrolysis is known as ripening, and the xanthate solution must be spun when it has progressed to a proper stage.

Suitable sources of cellulose are cotton linters, soda-boiled and bleached, or bleached sulfite pulp (p. 334), high in α-cellulose content.* The purified cellulose is treated with a

---

* Alpha-cellulose is that portion insoluble in 17.5% caustic soda at 20.0° C. Of the cellulose dissolved, the portion precipitated by neutralization with acid is called beta-, and that still remaining, gamma-cellulose. The fundamental cause of the differences in alkali solubility is unknown. The distinctions are presumably not a matter of length of

strong solution of caustic soda (18 to 20%) for about two hours, in an operation known as "steeping," conducted in a press by means of which the excess caustic is afterwards removed from the fibers. The resulting alkali cellulose (p. 489) is disintegrated to a crumbly, fibrous mass, in water-cooled kneading machines under careful temperature control. It will be seen later that it is desirable to have the viscosity of the viscose solution as low as possible commensurate with the highest strength of the product. It is the practice, therefore, to allow the disintegrated alkali cellulose, having the approximate composition $C_6H_{10}O_5 . 2$ NaOH, to stand in drums or in pans in a room the temperature of which is carefully regulated. During this "ageing" or maturation process, the cellulose is degraded and consequently on subsequent solution either as viscose or in cuprammonium hydroxide (p. 154) a corresponding reduction in viscosity is obtained (Fig. 2). Oxidation occurs, but must be carefully limited. Indeed, oxidizing agents such as peroxides have been employed but now meet with little favor. The assumption is that degradation is a question of oxycellulose formation, although the degree of oxidation necessary is probably extremely small, since one break in the middle of the cellulose chain can result in a very great

FIG. 2. Changes during Ageing.*

the molecular chain; at least, they do not parallel copper ammonium viscosity. Under-cooking can yield a product of low α-content due to residual non-cellulosic, alkali-soluble impurities. Indeed, the only pulps of extremely high copper-ammonium viscosities known to the authors appear to be in this class. Unsatisfactory conditions of bleaching can give pulps of substantially equal viscosities which vary in α-content from 65 to 95 %, in which the alkali-soluble fraction is practically proportional to the carboxyl content. This is presumably an oxidation effect. Apparently low chain length itself can impart alkali-solubility to cellulose, but only at low viscosities. Still other factors may play a part.

* HEUSER, E., and SCHUSTER, M., *Cellulosechemie* 7, 17 (1926). Viscose-Viscosity Curve from PITTER, A. V., *Jour. Soc. Dyers and Colourists* 51, 21 (1935).

decrease in chain length.   The treatment results in a solution of lower viscosity for the final spinning process.

After ageing, the alkali cellulose is placed in rotating drums or churns and about 30 to 40% of carbon disulfide, based on the original weight of dry cellulose,* is added through a spray-pipe.   Xanthation probably occurs at random along the cellulose chains and usually requires from two to four hours, evolving considerable heat and necessitating careful temperature control.

The light orange, crumbly cellulose xanthate is dissolved in sufficient dilute caustic soda (roughly 3%) to give a viscose solution containing from 7 to 8% cellulose and about 8% free caustic soda.   At this point the average length of the cellulose chains is of the order of 400 to 500 glucose units (see p. 151).   As stated above, this solution is unstable and the cellulose xanthate undergoes a slow hydrolysis of the xanthate radicals, yielding products with progressively lower ratios of xanthate to cellulose. These reactions are accompanied by marked changes in viscosity of the solution, shown diagrammatically in Fig. 3, the viscosity progressively decreasing to a minimum.   The subsequent rise finally becomes abrupt, the solution eventually setting to a solid gel.

FIG. 3.   Ripening of Viscose.†

In the freshly formed xanthate, almost all the carbon disulfide is present as xanthate radicals combined with the cellulose.   As "ripening" progresses, there is a steady decrease in xanthate sulfur which can be measured, e.g., by precipitation of the cellulose with alcohol.   As time goes on, the sulfur accompanying this precipitated cellulose goes down to 15 to 25% of the initial value.   Meanwhile, if at any stage in the ripening the cellulose be regenerated, purified, and tested by solution in

* One cellulose unit ($C_6H_{10}O_5$) to one of carbon disulfide would require 47%.
† PARK, R. H., Mass. Inst. Tech. Thesis (1938).

cuprammonium hydroxide solution, it is found that there is no substantial change in its cuprammonium viscosity and hence in its chain length. It seems likely that the initial decrease in viscosity during the early stages of ripening is due to a progressive dehydration of the cellulose molecule consequent upon the loss of the presumably highly hydrated xanthate radical. This dehydration would result in a decrease in effective particle volume with a consequent decrease in viscosity. That desolvation of this sort can reduce the viscosity of an emulsoid has been shown elsewhere (see p. 172). Ultimately, the cellulose loses enough xanthate radicals to render it insoluble in water. It is well recognized that precipitation of an emulsoid can be preceded by a viscosity increase which may be sufficient to cause gelation (see pp. 185 and 243). This gelation is perhaps caused by an increase in effective chain length due to the bonding of adjacent chains by partial-valence forces.

Maintenance of the correct viscosity of the xanthate solution is extremely important in the spinning process. Low viscosity of the spinning solution increases the rate of flow through the spinnerets and the capacity of the spinning equipment. Even more important, however, is the maintenance of constancy of both viscosity and cellulose concentration of the spinning solution in order to get uniformity of the spinning operation. Consequently, the manufacturer carefully controls the ripening operation to secure proper viscosity avoiding, however, the steep rise at the right hand end of Fig. 2. This has the advantage that the hydrolysis has proceeded to a considerable degree before the spinning operation itself, so that the chemical precipitation or coagulation, which follows spinning and which is, in fact, a completion of this hydrolysis, can be accomplished with more dilute solutions, in less time and much more uniformly throughout the filaments than would otherwise be possible. In consequence, the control of the ripening operation is one of the most important keys to the success of viscose manufacture. Under ordinary conditions 40 to 50% of the xanthate originally formed is allowed to hydrolyze during ripening, requiring four or five days. To control the reaction rate, extreme care is necessary in maintaining constant composition and temperature.

During the later stages of ripening the solution is freed from suspended particles, *e.g.*, incompletely dissolved cellulose, etc., by careful filtration, after which it is de-aerated under vacuum, to prevent escape of bubbles of dissolved air in spinning. These operations can be so conducted as not to interfere with the ripening reactions.

In general, the viscosity of the viscose solution must be low enough to avoid mechanical difficulties in manipulation but, to secure a strong thread, degradation of the cellulose must be minimized. The viscosity depends first on the quality of the original cellulose, perhaps best estimated by the viscosity of the cellulose in cuprammonium solution. This quality is profoundly influenced by the preparation of the fiber; the process of manufacture of viscose pulp is controlled throughout to produce the maximum α-cellulose content and cuprammonium viscosity compatible with suitable reactivity towards caustic soda and carbon disulfide. Viscosity is also reduced by the degradation which accompanies ageing of the alkali cellulose (see p. 154). The final control of viscosity, which is based on the amount of hydrolysis during ripening, does not directly affect the quality of the final product, but is sharply limited in its viscosity-reducing capacity and requires most careful control.

The formation of the actual fiber or filament is achieved by the spinning operation, consisting essentially of forcing the viscose solution through fine orifices under high pressure into a coagulating bath. The special alloy pumps for this purpose are of the gear rather than the piston type, to minimize pulsations. It is usual practice to spin a number of filaments simultaneously by the use of a spinning jet or spinneret containing from 24 to 50 orifices, about 0.1 mm. in diameter. The jets may be either glass, or, more usually, acid-resisting metals or alloys such as platinum, platinum-iridium, or tantalum. The solution, exuding from the spinneret at the rate of 50 to 150 meters per minute, is coagulated by a bath of dilute sulfuric acid and sodium sulfate (or other addition agents) which completes the hydrolysis of the xanthate and regenerates the cellulose. The necessary concentration and distance of travel in the acid bath depend upon the composition of the viscose solution and the size of the fiber.

The group of filaments from each spinneret is twisted together into a thread after it leaves the bath, by an arrangement such as that in Fig. 4, where the thread is drawn from the bath over the so-called Godet pulley D, and collected in the box F, which is rapidly rotating. The twist is imparted by having the peripheral speed of the pulley and the angular rotation of the box in the correct ratio, the usual twist being 2 1/2 or 3 turns per in. The thread is coiled in a compact cake in the spinning box by its own centrifugal force.* It is also possible to spin the thread onto a bobbin, which, however, does not produce quite as good results as the centrifugal collector described above. Spinning under tension, accompanied by greater elongation of the still plastic thread, can greatly increase tensile strength, and advantage is being taken of this, particularly in the cuprammonium process. From the cake, the thread, still containing some soluble salts and an appreciable amount of elementary sulfur resulting from the decomposition of the various sulfur compounds, is unwound directly into skeins. The impurities are removed by washing and treatment with dilute sodium sulfide. The thread may also be bleached and sized, and is finally dried to a moisture content of about 10 to 20%. A flow sheet of a typical sequence of operations is shown in Fig. 5.

Viscose thread has a high luster,† greater than that of pure silk. For many purposes this is undesirable and delustering agents, such as oils, waxes, and the like, are dispersed in the

FIG. 4. Centrifugal Spinning of Viscose (Diagrammatic).

A. Spinneret.
B. Coagulating Bath.
C. Guide Rod.
D. Godet Wheel.
E. Guide Tube.
F. Accumulated Thread.

---

* The tube E is moved up and down by a mechanism not shown, thus distributing the thread uniformly in F. The constrictions in E prevent the thread from sticking to its sides.
† Which can be increased by spinning under tension.

spinning bath. These are held by the fiber and serve as centers for diffuse reflection. Even better results are obtained by dispersing titanium dioxide in the viscose solution before spinning, the particle size being so adjusted that the pigment is not removed in the filtration process. Numerous other materials have
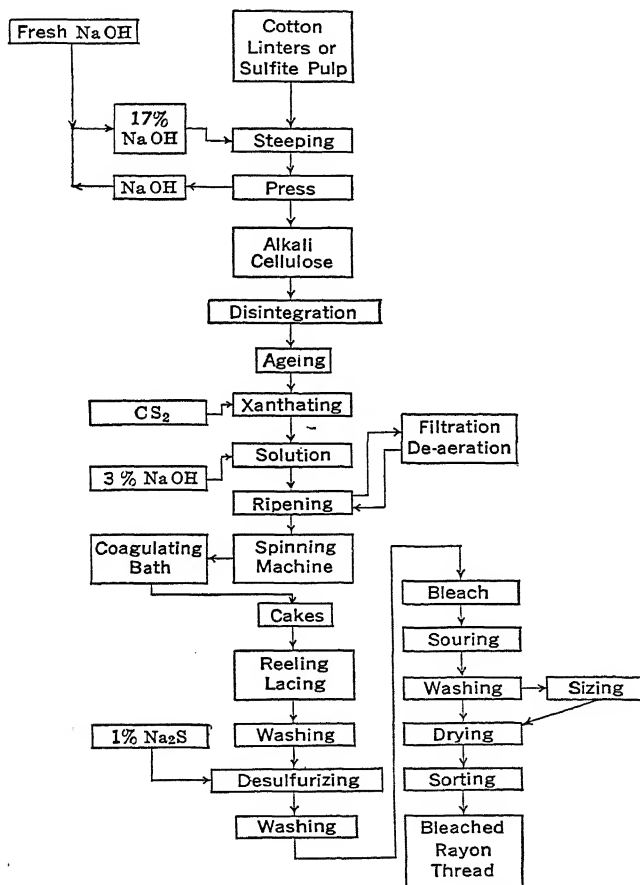
Fig. 5. Flow Sheet for Viscose Process.

been suggested, but all serve as centers for diffuse reflection, breaking up the continuous specular reflection of the viscose fiber.

An important application of the viscose process is the production of sheets and films of transparent, regenerated cellulose, called Cellophane. This material is formed by extruding the

viscose solution through a slot of the proper dimensions into a coagulating bath. The coagulated film is then washed, desulfurized, and dried in the usual way. A small amount of glycerol is often incorporated to improve luster and flexibility of the finished sheet. By giving the sheet a coating of cellulose ester lacquer a moisture-resistant product is made.

Cuprammonium Process.

Cellulose is soluble in solutions of copper oxide in ammonia, *i.e.*, cuprammonium hydroxide (Schweitzer's reagent). The
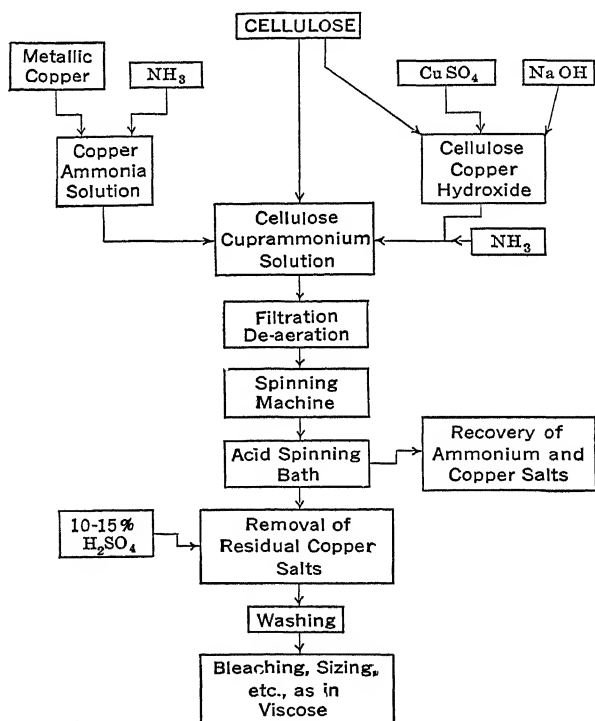
FIG. 6. Flow Sheet for Cuprammonium Process.

precise mechanism of solution is unknown; possibly a complex co-ordination compound is formed.* The viscosity of cellulose

---

* While it has been suggested that the cellulose takes no part in any chemical reaction, being simply dispersed by solution of the cementing pectic material, which holds the individual cellulose particles or micelles together in the original fiber, this seems incompatible with the fact that cellulose can be regenerated repeatedly from cuprammonium

in cuprammonium solution is a valuable guide to its degree of degradation; thus, undegraded α-cellulose shows high viscosity, undoubtedly due to long molecular chains (p. 166). The remarkable correlation between cuprammonium viscosity and other physical and chemical properties of cellulose indicates no substantial change in chain length on solution in Schweitzer's reagent. At one time a relatively large amount of rayon was made by the cuprammonium process and a superlative product may be produced, in particular, extremely fine filaments as low as 1 denier,* with resultant desirable qualities of soft handle and luster. The technique of spinning is essentially that used in the viscose process. A flow sheet is shown in Fig. 6.

Nitrocellulose.

Cellulose nitrate, although the first source of artificial silk, is no longer important for this purpose, but still remains one of the most widely used plastics and the basis of a major class of explosives. All three of the hydroxyl groups of the cellulose radical are capable of reaction with nitric acid to form the ester (technically but erroneously called nitrocellulose). The following equation is expressed per cellulose unit (or glucopyranose ring):

$$
\begin{array}{c}
CH_2OH \\
CH-O \\
-O-CH \qquad CH- \;+\; 3\,HNO_3 \;\rightarrow
\end{array}
\qquad
\begin{array}{c}
CH_2O.NO_2 \\
CH-O \\
-O-CH \qquad CH- \;+\; 3\,H_2O \\
\end{array}
$$

CH—CH   CH—CH
| | | |
OH OH   O₂N.O O.NO₂

(CH₂OH / CH—O / -O—CH ... CH— + 3 HNO₃ → CH₂O.NO₂ / CH—O / -O—CH ... CH- + 3 H₂O ; CH—CH with OH OH → CH—CH with O₂N.O O.NO₂)
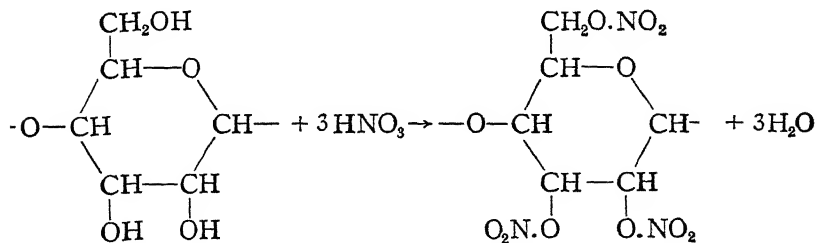
The reaction, like all esterifications, is reversible and is catalyzed by acids. Equilibrium favors nitrate formation only when the water content of the acid is relatively small, so that nitration requires concentrated acid. Since the reaction with nitric acid evolves water, the reacting acid dilutes itself, correspondingly

solution without substantial change in properties. It is difficult to see how any pectic bonding material between the original cellulose particles could be restored adequately on precipitation.
    * The metric denier is the weight in grams of 9000 meters of the thread.

slowing up the reaction and leading to an unfavorable equilibrium. Both effects can be counterbalanced by introduction into the mixture of sulfuric acid, and commercial nitrations are always conducted in mixed acid. The degree of nitration is controlled by the concentration of the acid employed, the mol fraction of water being the major factor. The temperature is kept down to prevent as far as possible side reactions, such as degradation by hydrolysis and oxidation. High acid concentration can minimize degradation, evidently because of lowered hydrogen-ion concentration, with corresponding reduction of hydrolytic rupture of the cellulose chains. The complexity of the relationships is indicated in Fig. 7.
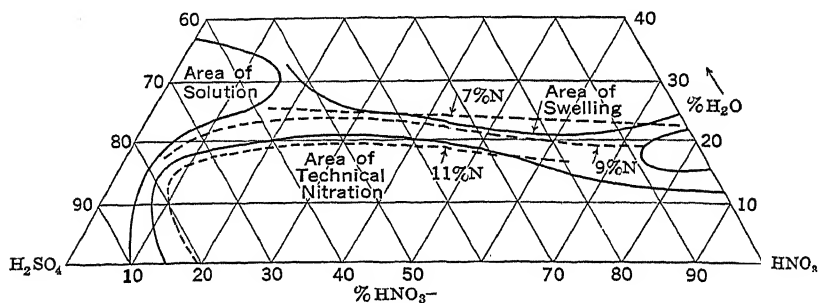
FIG. 7. Nitration of Cellulose. Dotted lines are for esters of equal nitrogen content. Within the area bounded by the full curves the fiber structure of the cellulose is destroyed.*

The properties of the resultant nitrate depend not only upon the extent to which the hydroxyl group has been replaced, but also upon as yet poorly defined reactions or changes in structure, which are determined by the strength of the mixed acids and by the time and temperature of reaction. Thus, traces of sulfuric and nitrosylsulfuric esters undoubtedly form and nitric acid exerts an oxidizing action on the cellulose, while both acids catalyze the hydrolysis of both the cellulose itself and its esters.

Complete replacement of the hydroxyl groups, yielding the trinitrate as indicated above, should give a product containing 14.15% nitrogen. In practice, this figure is never reached, although in the manufacture of smokeless powder products above

* MILES, F. D., *Trans. Faraday Soc.* **29,** 110 (1933).

13.9% nitrogen are produced. For ordinary commercial uses, the conditions of acid concentration, time, and temperature are adjusted to yield products from 10.5% up.* The structure of incompletely nitrated cellulose is a matter of debate. X-ray diffraction patterns have been claimed to indicate mixtures of completely nitrated trinitate with cellulose which is un-nitrated or relatively so. This conclusion seems improbable in the light of the solubility characteristics of the various products discussed below and of kinetic considerations,† particularly as solvent fractionation of a given nitrated product shows little or no variation in nitrogen content but wide differences in viscosity.

From the point of view of their properties and commercial utilization, the nitrates may be considered as belonging to three groups. The first contains approximately 11% nitrogen,‡ is very soluble in both a 1 : 2 alcohol-ether mixture and absolute alcohol, and is used for the manufacture of celluloid and related plastics. The second, containing approximately 11.5 to 12.3% nitrogen, is equally soluble in the alcohol-ether mixture, but somewhat less so in absolute alcohol, and is employed for lacquers, photographic film, and artificial silk. Nitrate containing 13% nitrogen or over is used for smokeless powder. It is substantially insoluble in both the solvents mentioned. All three types are soluble in acetone and the esters of the lower aliphatic alcohols, e.g., butyl acetate. These solubility relations of cellulose nitrates of varying nitrogen content are not sharp but the trends are indicated in Fig. 8, in which, however, each curve represents an average of a band of varying width.

The success of the commercial process of nitration demands uniformity of reaction. Furthermore, to reduce the degradation of the cellulose, it is advisable to nitrate the cotton as rapidly and at as low a temperature as possible. To obtain these con-

---

* Nitrocellulose for use in dynamite has a nitrogen content of about 12.5% and a chain length of 3000 to 3500 glucose residues; for use in plastics, the corresponding figures are 10.5 to 11.5 and 500 to 600; and for lacquers, 11.5 to 12 and 150 to 200.

† It is not impossible, however, that in a given glucopyranose radical the introduction of a single nitrate group renders the two remaining hydroxyls sufficiently sensitive to reaction so that complete replacement occurs in that radical. This would visualize a partly nitrated cellulose molecule as consisting of a chain of certain glucopyranose radicals completely nitrated and others unattacked.

‡ Varying in nitrogen content in special cases from 10.7 to 11.2%. The percentages given should not be taken as absolute but as indicative of usual ranges.

ditions all commercial nitration processes use a large excess of mixed acids (at least 30 to 1). Were a small amount of acid used, the consumption of nitric acid and generation of water would bring about a large change in acid concentration. Since the final acid must be at least as strong as that corresponding to equilibrium with the nitrate desired, the initial acid would tend to local over-nitration. While theoretically this could be corrected by saponification occurring in the later stages of the reaction, it obviously tends toward non-uniformity of product. The use of a large ratio of mixed acid to product makes it possible to start with a very low excess acid concentration so
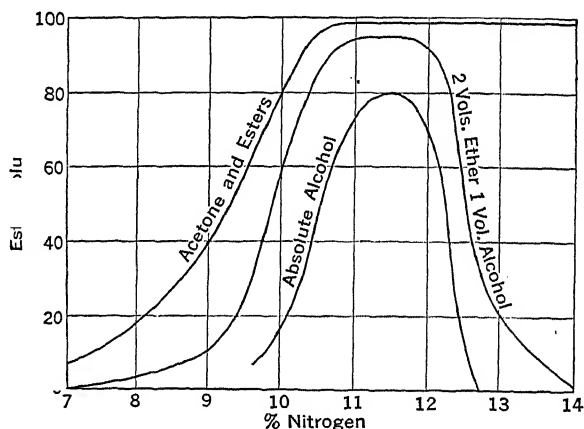
FIG. 8. Solubility Trends of Nitrocellulose.

that at no time or place is there a tendency for significant over-nitration. The acid initially penetrating to the center of the fiber is obviously weak and time must be given for diffusion to equalize acid concentration throughout the fiber. Even in the middle of the fiber, however, the tendency is always towards a progressively increasing degree of nitration.

The older method of producing these results consisted in immersing small batches of cotton in a large excess of acid, removing the cotton, and "drowning" it in a large excess of water. This dipping process, as it was called, was done entirely by hand and was therefore slow. Because of the fumes given off during the process, it was injurious to the health of the

workman, besides being extremely wasteful of acids due to the large amount retained by the cotton after nitration.

Today the dipping process has been profoundly modified by various mechanical methods of removing the spent acid from the cotton. Perhaps the most important technique is the removal of excess spent acid after nitration by means of a centrifuge, requiring only a few minutes, the acid thus slung out of the cotton being recycled to the nitrating bath after fortification with make-up mixed acid. The resulting cotton retains about its own weight of acid and is discharged to a drowning tank, often located on the floor below.

Another method is to immerse the cotton in acid contained in a wide shallow pan fitted with a perforated false bottom. A perforated cover is placed on the charge and a layer of water allowed to flow on the cover. If this is conducted carefully there will be no appreciable mixing of the acid and water layer. The nitration proceeds uniformly, the water layer serving to absorb any fumes that might otherwise be given off. On the completion of nitration the acid is drawn off the bottom of the pan and an equal amount of water allowed to flow in at the top. The water displaces the acid in the cotton with almost no intermixing, so that most of the strong acid is recovered. This process has the advantage of simplicity, low cost of maintenance of equipment, and low acid loss. However, it requires more labor and is considerably slower than the centrifugal process.

The nitrated cotton is carefully washed. Esters of the sulfate and nitrosulfate type, causing instability of the nitrate itself, are hydrolyzed readily by weak mineral acids and are removed by boiling with very dilute acid solutions, followed by thorough washing with water. The washing may be conducted, at least in part, in a beater, where the fibers are disintegrated, facilitating both the washing itself and later manipulation.

In the manufacture of artificial silk the nitrocellulose was dissolved in alcohol-ether mixtures and either spun into a coagulating bath of water or hydrocarbon or directly into air and the solvent evaporated. However, due to the inflammability of the cellulose nitrate, it was necessary to remove the nitrate groups from the cellulose. This was accomplished after the thread had been spun by treating it with warm, dilute solutions of sodium and calcium hydrogen sulfides, which reduce the nitrogen content to a fraction of a per cent. Although very fine

counts can be spun, the higher cost of manufacture has made
the nitrocellulose process obsolete. A flow sheet is given
in Fig. 9.
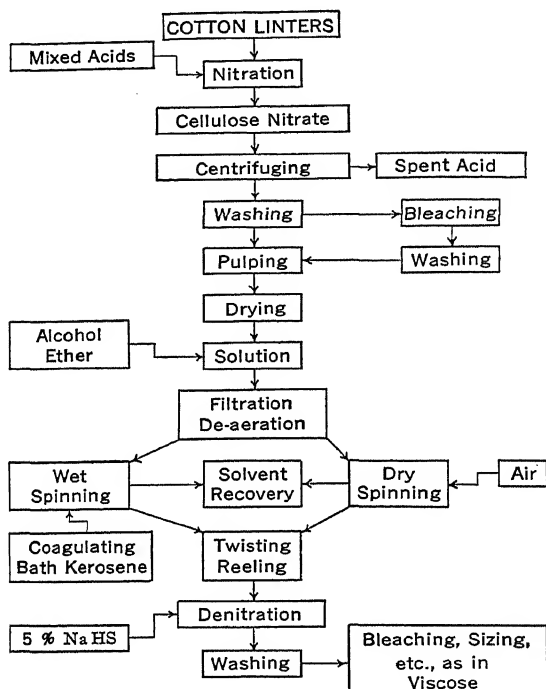
Fɪɢ. 9.   Flow Sheet for Nitrocellulose Process.

Celluloid : Pyroxylin.

Nitrocellulose is often used to produce molded plastics, the
ester being dissolved in alcohol with the addition of a plasticizer,
preferably camphor, the mixture molded into the desired form,
and the solvent evaporated.   A cellulose nitrate with a nitrogen
content of about 11.0% is used; this gives plastics with a greater
toughness than esters of higher nitrogen content.   The nitro-
cellulose is kneaded in a mill with the solvent and plasticizer
until a homogeneous, dough-like mass is obtained and much of
the solvent has evaporated.   The mix is then molded, or is
extruded through dies, while still containing about 10% solvent,
which is subsequently removed by storing at a slightly elevated
temperature.

Cellulose Acetate.

The chief disadvantage of viscose rayon is its marked sensitivity to water, the tensile strength of the wet fiber often being less than 50% of its value when dry.   This disadvantage has been overcome to a certain extent by the use of cellulose acetate as the fiber-forming material.

In the production of cellulose acetate, the cellulose, e.g., loose cotton linters which have been given the usual purification, is given a pretreatment to increase reactivity, as by steeping in glacial acetic acid containing catalyst * and allowing to age or mature, the time and temperature being carefully controlled as in the viscose process.   The pulp is then mixed with acetic anhydride in the presence of a catalyst, the most commonly employed being sulfuric acid, although materials such as zinc chloride, the oxychlorides of sulfur and phosphorus, etc. may be used.   In contradistinction to cellulose nitrate production it is unnecessary to use a large excess of acetylating reagent; 8 parts of 50-50 acetic anhydride-glacial acetic acid mixture per part of cellulose are sufficient.   The reaction is continued until the cellulose acetate dissolves in the mixture, producing a triacetate.
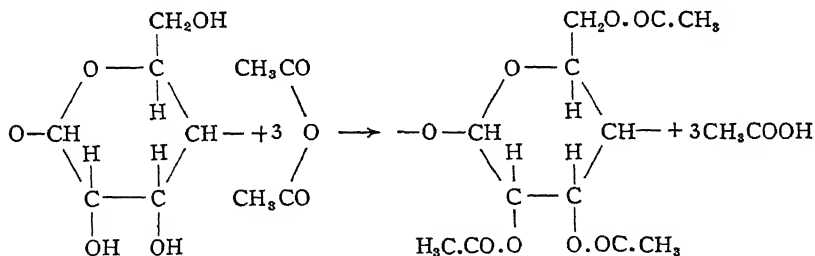
This compound (43 to 44% acetyl group), while soluble in chloroform, is insoluble in acetone and unsuitable for use in rayon, lacquers, or plastics.   It is, therefore, necessary partly to hydrolyze the acetate, by controlled dilution of the acetylating mixture with dilute acetic acid, to an acetate containing 38 to

* If cotton fiber be soaked initially in pure glacial acetic acid, it is rendered inactive, resistant to subsequent acetylation.   However, in the case of a wood cellulose, the character of which may be entirely satisfactory for acetylation, a preliminary soaking in pure acid prior to pretreatment in acid containing catalyst is essential to secure good activity of the cellulose in the final acetylation.   This anomaly is typical of the confusing relationships encountered again and again in the field of cellulose chemistry.

40% of acetyl groups.* Cellulose acetate for photographic film usually has a slightly higher acetyl content (39.5 to 42.0%), that for molded plastics a slightly lower one (36.5 to 38.0%). This partly hydrolyzed material is soluble in acetone but only swells in chloroform. After the hydrolysis, a large excess of water is added, precipitating the acetate, which is washed, centrifuged, and carefully dried at a low temperature. The product usually has a chain length of 175 to 350 glucose residues.

Fig. 10. Flow Sheet for Acetate Process.

It will be noted that large quantities of acetic anhydride are required for the reaction. Furthermore, the acetic acid must be recovered from dilute solution, as by fractional distillation or by solvent extraction with a suitable solvent such as a mixture of petroleum ether and ethyl ether. The weak acid washings may

* This is hydrolysis of the ester; hydrolysis of the cellulose chain does not occur.

be converted into sodium acetate, which can be recovered by evaporation. The high cost of acetic anhydride and the expense of recovering the acetic acid have in the past seriously limited the field of application of cellulose acetate.

The dry cellulose acetate is dissolved in acetone to give a clear, colorless, very viscous solution, usually of about 20% concentration, which, as in the viscose process, before spinning is first very carefully filtered to remove microscopic impurities and then de-aerated. Spinning may be accomplished either by forcing the solution through the spinneret into a precipitating bath, usually kerosene, or oftener by extruding the solution directly into warm air and evaporating the solvent under carefully controlled conditions. A flow sheet is shown in Fig. 10.

Besides the superior product obtained by the use of cellulose acetate, the dry spinning process is so simple that much higher speeds (200 meters per minute and over) can be run. Furthermore, the fiber requires no finishing treatment whatever after the spinning operation.

Cellulose Acetate Plastics.

Cellulose acetate is also suitable for the manufacture of molded plastics and film. Thus, it can be produced in sheets thin as 0.001 in. by dissolving in a suitable solvent, e.g., acetone, with plasticizer and allowing the solution to flow over a highly polished, revolving, metal drum. The solvent is removed by evaporation and a continuous film produced which is much less sensitive to moisture than the corresponding viscose film. Cellulose acetate is also thermoplastic at moderate temperatures, 280 to 340° F., and can be formed into any desired shape by compression, extrusion, or molding after admixture with a suitable plasticizer (see p. 319).

Properties of Artificial Silk.

The strength of rayon fibers, besides being dependent upon the cellulose chain length, is intimately connected with the degree of orientation of the cellulose chains in the extruded thread. Thus, ordinary rayon probably has a structure like that shown in Fig. 11(a), whereas a completely oriented fiber would appear as in Fig. 11(b). By spinning under tension, presumably induc-

ing orientation, very strong fibers can be obtained, the breaking strength being about three times as great as normal although the extensibility is much reduced. In the oriented fiber, all the chains reinforce each other and there is no tendency for flow after the yield value of the outer layer has been exceeded as in (a). Similar behavior is found with Cellophane, the tensile strength being greatest in the direction in which flow occurred on forming, and least in transverse sections across the sheet.

(a)

(b)
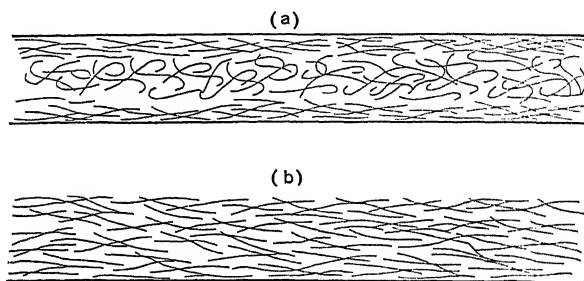
FIG. 11.

Apart from the behavior due to orientation, rayon fibers show differences in dry and wet strengths according to their mode of preparation (Table I, Ch. XXI). Although all except acetate rayon consist of regenerated cellulose, they differ from cotton in that they show a marked loss in strength when wet, a loss much greater than that of silk. Acetate rayon is outstanding because, although its dry strength is slightly less than that of viscose, it has a much smaller percentage loss in strength on wetting. This accounts to some extent for its increasing use.

Viscose, cuprammonium, and nitro rayons, being regenerated cellulose, may be readily dyed by any cotton dye, although direct dyes are usually employed to preserve their sheen. Acetate rayon, on the other hand, presented a special difficulty since a fiber of this composition had, previous to the last twenty years, not been met in the textile industry. Special methods for dyeing this fiber had, therefore, to be developed (p. 512).

The synthetic fibers of the past have been cellulose or derivatives of it, prepared from the natural fiber and depending upon its high molecular weight for the chain length of the product.

Indeed, the chain length of the original molecule of cellulose was shortened by the conditions involved in the processes of conversion. Attention has already been called to the success in recent years in the synthesis of long molecular chains. An example is the condensation product of acids such as adipic and bases such as cadaverine (p. 146). Synthetic fiber of this general type has been developed commercially and is now appearing on the market. A derivative of vinyl acetate, presumably a cross polymer with vinyl chloride, is likewise being introduced for a similar purpose. The two are known under the trade names of nylon and vinyon. There seems little doubt that these new synthetic fibers are not only likely to prove competitive with silk but are in many respects superior to it. The next few years will probably see an extensive transformation in the sources of textile fibers of the highest grade.

Esters and Ethers of Cellulose.

Although other esters of cellulose and such ethers as ethyl cellulose are not used at present in the form of artificial fibers, they are employed for plastics and lacquer compositions (p. 322) and, owing to their close kinship to cellulose acetate and nitrate, are conveniently considered here. Mixed esters, such as cellulose aceto-propionate and aceto-butyrate, are being manufactured by treating cellulose with mixed acid chlorides in the presence of a base (frequently pyridine). Unlike the acetate, no hydrolysis treatment is required. These mixed esters possess enhanced water resistance and a greater compatibility with solvents, gums, and plasticizers than cellulose acetate (p. 373). The ethers, another type of cellulose compound of promise, are prepared by heating alkali cellulose with alkyl chlorides or sulfates in suitable solvents, such as benzene, in the presence of excess caustic. To get complete conversion to tri-ethers, repeated treatment is necessary, because of the diluent effect of the water formed by reaction,

$$C_6H_{10}O_5 + 3\,NaOH + 3\,C_2H_5Cl \rightarrow C_6H_7O_2(OC_2H_5)_3 + 3\,NaCl + 3\,H_2O.$$
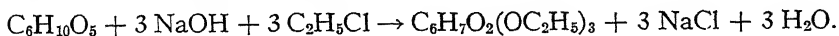
The ethers resemble the esters in properties, but possess unique solubility relations. Thus, ethylcellulose having 0.5 to 1.0 ethoxy group per glucose residue is soluble in dilute, aqueous

caustic soda (4 to 8%); with from 0.8 to 1.4 ethoxy groups it is water-soluble; above this, it becomes less and less soluble in polar solvents but retains its solubility in non-polar liquids. The ethoxy group does not usually confer water-solubility, but the explanation in this case is readily seen if the cellulose chains are pictured as bound together by the attractions of their hydroxyl groups, as described elsewhere (p. 349). The introduction of a few ethoxy groups tends to lessen the mutual attractions of the chains, thus promoting dispersion in water. Further substitution, however, will lead to a continuously diminishing number of hydroxyl groups and solubility in polar solvents will, therefore, diminish. Commercial ethylcellulose, with 47 to 48% ethoxy content, contains about 2.4 to 2.5 ethoxy groups per glucose radical and is soluble in most alcohols, ketones, esters, and hydrocarbons. For lacquer work it possesses high miscibility with plasticizers and resins, and is very resistant to light.

One must remember that the viscosity characteristics of all cellulose esters and ethers are governed by the degree of degradation of the cellulose. Thus ethylcellulose, like nitrocellulose, can be obtained with any desired viscosity by pretreatment of the cellulose before ethylation.

BIBLIOGRAPHY

BONNET, F., "Vinyon," *Ind. Eng. Chem.*, **32**, 1564 (1940).

BRANDENBURGER, J. E., "Cellophane," *J. Frank. Inst.*, **226**, 797 (1938).

DORÉE, C., *Methods of Cellulose Chemistry*, D. Van Nostrand (1933).

DURRANS, T. H., *Solvents*, D. Van Nostrand (1938).

HOFF, G. P., "Nylon as a Textile Fiber," *Ind. Eng. Chem.*, **32**, 1560 (1940).

MARSH, J. T., and WOOD, F. C., *An Introduction to the Chemistry of Cellulose*, D. Van Nostrand (1939).

PELLATT, D. L., *Viscose Rayon Production*, Emmott and Co., Ltd., London (1931).

SCHWARZ, E. W. K., and MAUERSBERGER, H. R., *Rayon and Staple Fiber Handbook*, Ed. 3, Rayon Handbook Co. (1939).

Development and Use of Rayon and Other Synthetic Fibers, Report of U. S. Dept. of Agriculture (1938).

WATSON, W., *Textile Design and Colour (Appendix)*, Longmans (1936).

WELTZIEN, W., *Kunstseide*, Akad. verlag. m.b.h., Leipzig (1930).

WHEELER, E., *Manufacture of Artificial Silk*, Chapman and Hall, Ltd., London (1931).

Chapter XVII

Leather

Gelatin (p. 307) and leather form the major commercial products derived from animal skins, which, in turn, are in large part a by-product from the butcher and packing-house industry. Before considering the production of leather, it is important to study the structure of the animal skin and the proteins comprising it.

Skin Structure.

Although skin structure varies greatly with species and the location in the individual skin, certain broad generalizations are possible. The skin is divided into two distinct parts, a thin outer layer, the epidermis, usually forming about 1% of the total thickness, and an inner layer of interwoven and connective tissue, the derma or corium. In skins as received by the tanner, a third layer may also be present, the fascia or fleshy material, which in reality belongs to the body of the animal. The outer layer or epidermis is a thin layer of bloodless cells which receive their nourishment from the blood vessels of the derma. As the cells reproduce, the older ones are pushed outwards, gradually die and become dry and scaly, easily rubbing off. At the bottoms of pockets or follicles the hairs grow through this layer. The "sebaceous," or oil glands, and the "sudoriferous," or sweat glands, drain through tubes which also penetrate the epidermis. The derma consists of a reticulated fabric of closely interwoven, and to some extent intergrown, fibers of the protein, collagen. Blood vessels, nerves, and fat cells are situated throughout the fiber network, while a bonding protein tissue, called elastin, also occurs, especially in the upper and lower surfaces of the derma. Below the epidermis, the corium is dense and chemically resistant, forming the basis of the desired "grain surface" of

leather.   Its structure may be seen from Figs. 2 and 3, pp. 390–
391.   The behavior of the varying layers of the skin toward
chemical reagents is of great importance, since in either gelatin
or leather manufacture the epidermis and fascia must be re-
moved and the corium left available for further treatment.

The epidermis appears to consist principally of the protein,
keratin, which, it will be seen later (p. 493), is unusual in
containing the cystine residue with its disulfide (—S=S—)
link.   This linkage is readily attacked by alkalies and sulfides.
In contrast to keratin, at room temperature collagen is more
susceptible to hydrolysis by acids than by bases.   Thus, it is
possible by suitable treatment with bases (p. 382), completely
to remove the epidermis, leaving the collagen comparatively
unattacked.   As already seen (p. 307), hot water causes collagen
slowly to go into solution as gelatin, elastin being unattacked.
On the other hand, the enzyme, trypsin, readily dissolves elastin
but leaves collagen and keratin unchanged.

Structure of Leather.

Skins, particularly of the mammalia, constitute a natural raw
material of extraordinary physical properties, a material which
becomes even more valuable when properly modified in struc-
ture.   As already pointed out, the animal skin consists of a
natural fabric of interwoven fibers, called rawhide, which, while
tough and strong, possesses two exceedingly serious limitations.
In the first place, although insoluble, because of its protein
structure it is extremely sensitive to water, swelling to a stretchy
mass and being somewhat plasticized.   On removal of the water
by evaporation the plasticized fibers cement themselves together
into a hard and horny solid.   In the second place, the wet skin
is extraordinarily sensitive to putrefaction.   Therefore, the skin
must be treated to reduce its sensitiveness to water and render
it resistant to putrefaction.   The product is called leather, and
the corresponding conversion of the skin, tanning.   This is
done by combining with the protein of the derma, by either
chemical reaction or physical adsorption, some material which
will greatly reduce its hydrophylic character and protect it
against putrefaction, with a minimum alteration in either the

physical properties or the interrelationships of the ultimate fibers of the skin. For this conversion a large number of materials are available, such as the tannins, the basic salts of various trivalent metals, formaldehyde and similar bodies, tungstic acid, etc. Of these, by far the most important are the vegetable tannins and chromic salts. Their use divides the industry into two major subdivisions, vegetable and chrome-tanning.

It is desirable for the tanner to be able to control the texture and feel of the finished leather just as the textile manufacturer can control the properties of a fabric. In doing this the tanner is faced by definite limitations. Whereas the character of the textile fabric can be modified, not only by the choice of the raw fiber but also by the weight and tightness of twist and weave, the tanner cannot correspondingly modify the structure of the original skin. However, he can loosen and soften the leather made from it by eliminating a part of the skin structure itself. Correspondingly, he can harden and stiffen the leather by filling the voids between the fibers and bonding the whole mass permanently together, as in the sizing of paper and textiles. The former is usually done to a greater or less degree in the manufacture of the flexible, so-called light leathers such as upper leathers for shoes, while the latter is characteristic of heavy leathers, *e.g.*, sole leather. For the former the chrome process of tanning is particularly adapted, and the vegetable for the latter.

Beam House Operation.

The skin may come to the tanner substantially as removed from the animal (green hide), or it may have been salted or dried to protect it from putrefaction during shipment. In any event it is contaminated with dirt and blood and usually has hair adhering, as well as the so-called trimmings, ears, tails, etc. For all initial operations in the tannery it is essential that the skin be thoroughly hydrated. In consequence, the first step in the tannery is soaking and washing the skins. Green hides do not require long soaking, but the hydration of dry hides is always slow. The operation is sometimes conducted in pits, particularly in the initial stages. The final washing is usually carried out by tumbling in large, rotating drums in a stream of water;

small amounts of water-soluble proteins are extracted. Trimmings are removed by hand. The flesh is removed by passing the hide over rolls beneath the rotating knives of the fleshing machine.* From the skins thus cleaned and softened the hair and epidermis are removed by selective chemical or bacterial action. The oldest method, sweating, took advantage of the fact that the tissue at the base of the hair follicles is more sensitive to putrefactive as well as chemical attack than the rest of the skin. The wet skins were hung in a warm, humid room until the hair was loosened.† Because of the irregular action, it was difficult to avoid injury to a certain percentage of the skins.

A method almost as old is liming—soaking the skins for a number of days in a dilute milk of lime at normal temperature. This attacks the epidermal tissue at the base of the hair follicle, the protein structure of which, probably keratin in an early stage of formation, is the most sensitive point in the whole skin under these conditions of alkalinity. Old lime liquors are much more active in their depilatory action than new and have been shown to contain amines as active auxiliary depilatories, evidently formed by some sort of protein decomposition. It is now generally considered that depilation is effected by chemical hydrolysis of the disulfide linkage,‡

$$R—S—S—R + H_2O \rightarrow R—SH + R—S—OH.$$
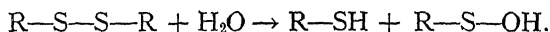
* This may be done after dehairing.

† It seems likely that the long-continued survival of the sweating operation was due to the fact that the putrefaction not only loosens the hair but destroys a small percentage of the fibers of the skin, thereby opening up its structure, to make it slightly more distensible for the production of heavy leather.

‡ The acid formed may break down,

$$CH.CH_2.S.OH \rightarrow CH.CHO + H_2S,$$

the product being again capable of cross-linking;

$$CH.CHO + H_2N.(CH_2)_2.CH \rightarrow CH.CH : N[CH_2]_2.CH + H_2O.$$
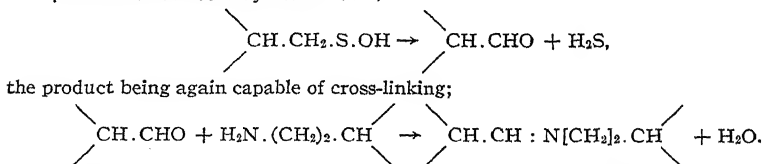
On the assumption that dehairing can only be accomplished by breaking all cross-linkages of any kind, it can be seen that fresh limes would be slow. Old limes, on the other hand, contain amines of low molecular weight which can condense with and eliminate the aldehydes as fast as formed. The question still remains as to why the amines are generated more rapidly in old limes than in fresh. It may be that their origin is in bacterial decomposition of the proteins, which will only occur at maximum rate after the productive organisms have established themselves; or, the proteins may break down at a fairly constant rate into an intermediate material which in turn breaks down only slowly into the amines themselves. These form at a reasonable rate only provided the intermediate material is allowed to accumulate to a relatively high concentration.

The most rapid technique is solution of the hair in dilute sodium sulfide, but it is debatable whether this method does not significantly affect the collagen because of the high alkalinity. However, the sulfide is frequently added in small amounts to the limes to "sharpen" their action.

Skins carrying wool which it is desired to salvage uninjured may be painted on the flesh side with a paste such as calcium hydroxide in sodium sulfide solution. This penetrates the skin and attacks the tissue at the base of the hair follicles so that the wool may be removed completely unaffected by the treatment.

Whatever method is used, the loosened hair and epidermis are readily removed by scraping with dull knives, usually set into the roller of the unhairing machine. Hair left by the machine is taken off manually by scudding, the skin being thrown over a wooden slab (beam) and scraped with a two-handled knife.

In order to obtain a smooth and cleaner grain surface and softer leather, the washed skins after scudding are often "bated," *i.e.*, treated with pancreatic enzymes in the presence of a buffer solution, usually ammonium chloride, accurately to control the pH. By this means certain proteins of the skin are dissolved and eliminated, while the collagen or main fibers are but little attacked. Bating is not carried out on all skins, but is considered essential for light leathers which, to a large degree, are sold upon their finished appearance.

Vegetable Tanning.

The original fibers of the skin may be regarded as forming a water-sensitive organized gel. In vegetable tanning the means which the tanner employs to transform this is to precipitate upon the fibers a colloid of electric charge opposite to their own. Like all proteins, the substance of the hide is amphoteric in character (p. 207); in acid solutions the protein becomes positively, in basic solutions negatively, charged.

$$
\begin{array}{ccccc}
\overset{-}{COO} \; \overset{+}{Na} & & COOH & & COOH \\
\diagup & & \diagup & & \diagup \\
R & \xleftarrow{\;NaOH\;} & R & \xrightarrow{\;HCl\;} & R \\
\diagdown & & \diagdown & & \diagdown \\
NH_2 & & NH_2 & & \overset{+}{NH_3} \; \overset{-}{Cl}
\end{array}
$$
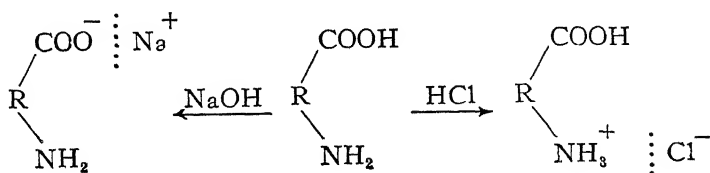
If brought into contact with any colloidal solution the particles of which carry a charge opposite in sign to that of the hide substance, co-precipitation (p. 130) results, producing a large aggregate. In this way not only are some of the hydrophilic groups blocked or rendered inactive, but the increase in size of the particle tends to render it more inert and more resistant to the action of solvents and reagents, chemical and otherwise (cf. p. 278). From this point of view there ought to be a large number of materials available for the conversion of hide substance into leather. This is indeed true, but experience has shown that most of them are, for one reason or another, unsatisfactory. Thus, for some reason positive colloids used as precipitants for hide substance in basic solution usually produce an unsatisfactory product. It is obvious from the description of the unhairing methods outlined above that all these are carried out in basic media. The tanner must, therefore, get his skin on the acid side before converting it into leather. When this has been done, he must still avoid any negatively charged colloids, the use of which might prove disadvantageous. The oldest and most satisfactory precipitants are the so-called tannins, mixtures of glucosidic condensation products of polyphenolic bodies, widely distributed in the vegetable kingdom.

Tannin solutions are obtained by countercurrent extraction of the wood, bark, nuts, or leaves of trees. Where the extract is to be transported any considerable distance, this is followed by vacuum evaporation, avoiding oxidation. They may be grouped into the astringent, powerfully precipitating tannins, such as quebracho, hemlock, and oak, and the non-astringent, mild, slightly but none the less effectively adsorbed tannins, such as gambier and sumac.

The capacity of the skin for taking up tannins is very great indeed. Thus, in the tanning operation, increases in weight of over 300% may be encountered. It seems probable that after enough astringent tannin has been absorbed completely to coagulate the protein of the fiber, the coagulated aggregate still has capacity to adsorb large additional quantities of the tanning agent, building up a particle of far greater size than that of the original fiber, thereby filling the voids and stiffening the

structure. In consequence, thick skins are tanned predominantly with the vegetable tannins to produce heavy leather. The two most important types are sole leather, in which resistance to abrasive wear and low flexibility are secured even at the expense of tensile strength, and belting, in which flexibility and tensile strength are the important considerations.

In order to understand the tanning operations certain important chemical factors must be kept in mind. The tannins are weak acids and behave as negatively charged colloids, being precipitated more readily by trivalent than by di- or monovalent ions. Although tannic acid will precipitate hide substance in a mildly alkaline medium, there is formed in the presence of lime a certain amount of calcium tannate, which is subject to oxidation and darkening by the oxygen of the air. Furthermore, the alkaline tanned material has poor wear resistance. It is, therefore, necessary to neutralize and wash out the lime or other alkali used in dehairing. However, as the hide is brought by this neutralization nearer to its isoelectric point it shrinks greatly (p. 211), and, if converted into leather in this condition, the product has essentially the thickness of the hide at the beginning of the tanning operation. By absorption of acid, however, the hide swells on the acid side of its isoelectric point, and if the hide be tanned while in this swollen condition and the pores adequately plugged up by precipitation of sufficient excess tannin, the finished leather will retain a thickness characteristic of the acid-plumped hide. Furthermore, quite a little incipient tanning by mild, non-astringent tannins may be allowed to occur without having the hide lose its capacity to swell in dilute acid. Now the natural tannins contain sugars, either present as such or produced by hydrolysis during the tanning process. These sugars can undergo fermentation during the tanning operation, producing considerable quantities of weak organic acids which accumulate in the spent tan liquor. These are, therefore, available for neutralization of the lime and plumping of the skin, although synthetic acids, either mineral or organic, may be used, either for pretreatment of the skin before any tanning at all is undertaken or for addition to spent tan liquors to build up their natural acid content. Finally, if one treat fresh hide with a

strong solution of astringent tannins, the precipitation on the surface of the hide is so violent that the surface pores are plugged up and further penetration of the tannin by diffusion through the hide is rendered impossible, or at least inordinately slow. It is, therefore, clear that the treatment of the hide by the tan liquor should be a countercurrent operation, consisting of a relatively large number of steps.

In consequence of these facts the standard procedure for tanning sole leather is to wash the skin from the beam house to remove most of the alkali and then treat with acid to neutralize and wash out the remaining lime and plump the skin. As indicated above, this may be done in connection with the first stages of the countercurrent tanning process, by treating the washed hides with spent tan liquors of proper acidity. The hides are then treated with stronger and stronger tanning liquids in a series of operations. In the first stages of the series the liquors are almost always mildly agitated, the hides hanging in them suspended from sticks or rocker arms or equivalent devices. Toward the end of the series the hides may be thrown on top of each other in pits, the so-called layers of the tannery. However, especially during the earlier stages of treatment, the stock is protected from violent agitation or flexing because this is likely to cause the leather to fall, *i.e.*, decrease in thickness. It would seem that after the tanning operation has gone to a certain extent, the affinity of the hide for the acid is largely lost so that the plumping action of that acid disappears. However, the hide will still retain its swollen shape unless handled roughly. Once the voids between the fibers have been plugged up with tannin, shrinkage becomes impossible. This whole operation involves the diffusion of material of extremely high molecular weight or particle size and, therefore, low diffusional velocity through the pores of the hide into the interior. Hence, it is a slow operation and weeks and even months are necessary to complete it. Even so, it is clear that at the end of the operation the voids are filled with a tannin solution which, though strong, still contains a large amount of water. It therefore becomes imperative to fill these voids with solids. This is done by dehydrating the hide as completely as practicable, *e.g.*, by

running through squeeze rolls, and then immersing it in a very concentrated tannin extract. This extract is imbibed by capillarity into the voids from which the water was removed. To give time for the maximum possible penetration of tannin into the leather, it is usually allowed to soak in a pit covered with warm, concentrated extract, an operation called *tempering*. Other inert loading materials may be worked in by similar methods.

It is clear that the large amount of tannin taken up by the skin is forced from the outside inward. Consequently, the heaviest loading is on the surface. The outer layer of the skin, the so-called grain surface, is built up of the most closely packed fibers and possesses a structure of greatest density. When one has submitted the skin as a whole to this loading operation to a degree sufficient adequately to plug up the voids in the center of the hide, the plugging effect on the grain surface has become excessive. The surface has been converted into a solid wall in which every bit of space between the fibers is filled up with dried-out, solid tannin. Obviously, the dried grain surface cannot flex. If a piece of sole leather in this condition be bent with the grain convex, it will always crack. If, on the other hand, the tanning operation be carried only to the point where the grain surface is properly loaded, the interior of the hide will be inadequately filled. The leather will be soft and flabby and, due to its capillary structure, will absorb water like a sponge. It is, therefore, imperative to tan the leather so that it is adequately loaded through and through, and thereafter remove the excess tanning material from the grain surface.

When the loading with tannin is completed, the leather is dried. This is done slowly to minimize the movement of tannin from the interior toward the surface by capillarity. This drying is apparently effective in dehydrating the tannins to a distinct degree, with the result that they are thoroughly fixed on the fibers. To remove the excess of tannin from the surface layer, the leather is dipped into warm water to fill its voids and then for a moment into warm, dilute alkali, usually sodium carbonate or borax. The alkali chemically combines with the acid tannins in the surface voids and takes them into solution. The skin is

then dipped in water to wash off these dissolved tannins, followed by dilute acid to neutralize the alkali, and finally by water to remove the surface acid. This operation is called *bleaching* because it removes the dark-colored, astringent tannins left in the surface at the end of the tanning operation, and the action of the acid is to lighten still further the color of those which remain. However, its most important function is the elimination of brittle grain by the removal of the excess tannins in the voids of the grain surface, thereby making it possible for the grain fibers to flex and bend.

Before the final drying of the leather a certain amount of oil or grease is introduced. This may be done by fat-liquoring, *i.e.*, drumming the leather in a water emulsion of the oil or fat; or by stuffing, *i.e.*, drumming in liquid fat, using heat if necessary; or by swabbing or oiling off, with clear oil or an emulsion. However, in the case of sole leather an oil emulsion (*e.g.*, sulfonated cod oil) is usually introduced with the tanning extract in the loading operation, additional fat being applied later by swabbing.

The final stages of finishing the leather involve moistening the dried product sufficiently to make it flexible, oiling off with emulsions part of which penetrate and part remain on the surface to give it a proper gloss, followed by rolling, to increase the density of the leather and give it a surface gloss, and finally drying to a moisture content in substantial equilibrium with the air.

The manufacture of belting leather differs from that of sole leather in three essential particulars. Belting must be strong and flexible, but not stretchy. Therefore, in the first place, in its manufacture one uses enough acid to remove the lime but not enough to plump the hide. This leaves the maximum amount of hide fiber per unit of cross section to give it the strength required. In the second place, while one tans thoroughly, one does not load to the limit. This gives enough filling of the voids to reduce stretch, but not enough to make the product stiff. Finally, belting is stuffed, *i.e.*, tumbled or dipped in grease (tallow and cod liver oil) warm enough to be fluid. This incorporated grease separates every fiber in the final product from

its neighbor and likewise furnishes the lubrication so necessary to reduce to the lowest possible point the friction between the fibers which results from their repeated flexing (see also p. 394).

Chrome-Tanning.

In contradistinction to heavy leather, one normally wishes in the manufacture of light leather to obtain a soft and flexible product, in which filling of the voids between the fibers has been avoided. This may be achieved, first, to a slight extent by modification of the treatment of the skins in the beam house *

FIG. 1. Finished Calf Chrome Tanned.† Frozen section; 25μ; no stain; mounted in glycerin jelly; $(45\times)$.

* It is possible to open up the skin here; thus, the use in the beam house of old limes (p. 382) results in the production of a softer leather, whereas dehairing with sodium sulfide or fresh limes gives a firmer product. Apparently, the old limes attack and disintegrate some fraction of the hide itself. Bating also removes specific structures, which affect the appearance of the grain surface of the finished leather, but probably in addition attack simultaneously some of the finer fibers of the main body of the skin.

† Courtesy of M. C. McDonald, A. C. Lawrence Leather Company, Peabody, Mass.

and, second, by greatly reducing the amount of tanning agent employed, eliminating filling of the interfibrous voids by tanning material, and minimizing any distention of the fibers due to tannin actually combined with them. While this is often accomplished by the use of mild, non-astringent vegetable tannins of limited combining tendency, the most effective technique is the process of chrome tannage, for reasons that will appear below.

In chrome-tanning the effective tanning agent is a complex basic chromium salt, which is soluble, even colloidally, only in solutions of low pH. The chromium is taken up by the hide

FIG. 2.   Bark Tanned Cow Hide.*   Paraffin section; 25μ; mounted in balsam; $(10\times)$.

substance from solution, the surprising fact being the small quantity of chromium required to give adequate tannage (2 to 2.5% chromium on the weight of the hide being capable of giving a satisfactory product), as compared with the large amount of tannin necessary in the vegetable-tanning process. This explains why leather tanned by the chrome process may be per-

* Courtesy of M. C. McDonald, A. C. Lawrence Leather Company, Peabody, Mass

fectly resistant to water and putrefaction, yet remain soft and flexible because the voids have not been filled in any way. Figs. 1, 2, and 3 show vertical sections through two similar portions of skin, one of which has been chrome- and the others vegetable-tanned. The difference is striking, the chrome-tanned skin appearing empty in comparison with the swollen vegetable-tanned skin, in which the interfibrillary spaces appear to have been almost completely filled.

FIG. 3.   Bark Tanned Cow Hide.*   Paraffin section; 25μ; mounted in balsam; (45 $\times$).

The mechanism of chrome-tanning is still uncertain; since the chromium complex under the usual conditions of tanning may carry the same charge as the hide substance, ordinary co-precipitation cannot be the explanation.   It is not unlikely that the tanning is due to a co-ordination compound of the chromium complex and the collagen.   The situation is further complicated by the fact that, at constant pH, fixation of the chromium appears to follow the Freundlich adsorption isotherm (p. 90) with

* Courtesy of M. C. McDonald, A. C. Lawrence Leather Company, Peabody, Mass.

respect to the chromium concentration of the tanning solution. The degree of initial combination increases rapidly with pH. Indeed, the chromium is introduced at low pH to secure thorough penetration, and combination with the skin induced by careful neutralization.

Whatever the reaction mechanism, when the chromium has been precipitated from aqueous solution on the skin the superficial properties of the hide have been changed to a remarkably slight degree. The slipperiness of the skin has been reduced decidedly but otherwise, except for the green color, its appearance is much the same. However, an extraordinary and far more fundamental transformation has occurred. Prior to the precipitation, the fibers possess a positive affinity for water. Furthermore, their reaction towards water is reversible. The skin can be dried out and wet back at will. At the time the chromium was added the skin was highly hydrated. After the addition the skin is still hydrated, but it is now in a condition where, once dehydrated again, it permanently loses its affinity for water. The skin is also resistant (unlike vegetable-tanned leather) to boiling water. In other words, the skin is now an irreversibly hydrated material.*

Since the chromium is introduced in acid solution, the skins must be highly acid before the process starts. Furthermore, because one wants a product which is soft and strong rather than stiff and thick, one must avoid swelling of the skin, which is essential in sole leather tanning. One must avoid it despite the fact that in preparation for the tanning operation the skin must be made far more acid than in the case of sole leather. Consequently, it is imperative to employ along with the acid some sort of agent which will prevent its plumping action. It has been pointed out that emulsoids can often be precipitated or their swelling reduced by the addition to the solvent of material miscible with the solvent but in which the emulsoid is

---

* The parallelism to the behavior of silicic acid in the formation of silica gel is striking (see pp. 183–186 and 238–240). Apparently, the fiber of the hide is transformed by the precipitation of the chromium in a similar way. The precipitation does not eliminate the water. The hide fiber is still quasi-plastic. If, however, one eliminates the moisture by drying, the rehydration of the skin, at least to anything approximating its former state, is no longer possible.

insoluble (p. 174).* The cheapest dehydrating agent of this sort available for the use of the tanner is common salt. Therefore, in the preparation of the hide for the process of chrome-tanning it is pickled, after washing or bating, in a solution of dilute sulfuric acid and salt, with relatively little of the former and large quantities of the latter. In this way one can bring the hide to the required acidity without excessive swelling. The distribution of the electrolytes between the skin and the external solution is governed by the Donnan equilibrium (p. 212).

The chrome liquor is usually prepared by the reduction of sodium dichromate solution, acidified with dilute sulfuric acid, by sulfur dioxide, glucose, molasses, or other suitable reducing agent. In this way a basic chromic

$$Na_2Cr_2O_7 + 3\,SO_2 + H_2O \rightarrow 2\,Cr.OH.SO_4 + Na_2SO_4$$
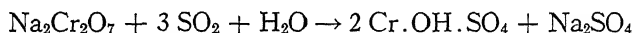
sulfate is obtained, which can be used directly for tanning. The pickled hide, steeped or rotated in a drum with the chromic salt, is penetrated through and through and if the acidity be not too high, considerable amounts of the chromium are actually adsorbed by the skin. In order to have it take up the chromium completely, partial neutralization of the acid is necessary. This should obviously be carried out gradually and progressively in order to get uniformity of action throughout the thickness of the skin and, in practice, it is effected by adding sodium bicarbonate or borax in successive small quantities.

After tanning, the goods are generally shaved and washed. Sometimes neutralization is not completed until this point. They are then ready for fat-liquoring (p. 388), imperative for chrome-tanned leather before drying.

Fat-liquoring emulsions usually consist of cod liver or neatsfoot oil, dispersed in water or in a solution of sodium carbonate or borax, with the aid of an emulsifying agent such as egg yolk or soap (see p. 252 for a discussion of emulsions). On diffusion into the leather the emulsion apparently breaks and the oil globules are deposited on the fibers. pH control is of great importance. Sulfonated oils are particularly useful because they

* For example, acetone will depress the swelling of rubber in benzene, alcohol that of gelatin in water, etc.

combine in large degree with the leather substance and are un-extractable with solvents. Unsulfonated oils, on the other hand, are readily extractable and appear to be mechanically dispersed through the leather. Any oil used must, of course, be non-injurious to the leather.

FIG. 4.   Patent Leather, Two Finish Coats.*   Paraffin section; 20μ; mounted in glycerin jelly; $(45\times)$.

It was formerly felt that the major function of fatty substances in the leather was waterproofing and lubrication of the fibers to promote flexing. No doubt these factors are important, but even more essential is the action of the fat in preventing coalescence of the fibers on drying. The tannins precipitated on and between the fibers are water-soluble. Furthermore, they form in water emulsoidal solutions which, when concentrated, become extremely viscous, sticky, glue-like materials. If, therefore, one dries down a piece of wet leather, as the tannin solutions between the fibers concentrate the fibers become glued together so that the whole mass sets and hardens as it dries to a veritable

* Courtesy of M. C. McDONALD, A. C. Lawrence Leather Company, Peabody, Mass.

piece of horn. If, now, in the final stages of manufacture, one incorporates into the leather globules of oil or fat, these will insulate the fibers from each other and to a large degree prevent their adhesion during drying. Therefore, in the dried products the fibers are relatively free to slip over each other, making it possible for the leather to flex. This is always a major function of the fatty substances incorporated in leather and sometimes their sole function. Thus, leather for japanning (*i.e.*, patent or
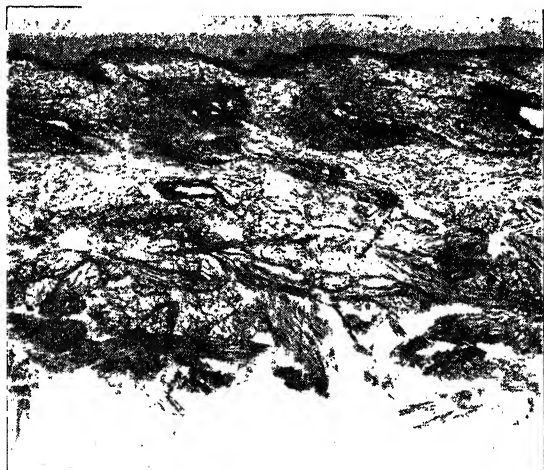
FIG. 5. Patent Leather, Three Finish Coats.* Paraffin section; 20μ; mounted in glycerin jelly; $(45\times)$.

varnished leather, Figs. 4 and 5) must be fat-liquored for this reason, but the grease incorporated must after drying be extracted by naphtha to prevent injury to the ultimate finish.

Chrome leather may be dyed either before or after drying. The dried leather is dampened back, tacked out on frames in a stretched condition, and again dried. The frames greatly reduce the shrinkage on drying, yielding a larger area of product and leaving a leather which will subsequently stretch far less when wet. The product can be finished by the application of a surface coat of pigments or dyes bonded together with agents such as blood albumin, casein, and the like, or surface dopes such as the

* Courtesy of M. C. McDonald, A. C. Lawrence Leather Company, Peabody, Mass.

drying oils or lacquer finishes of patent leather may be used. A gloss can be developed by pressing between smooth hot plates, or ironing or jacking by the friction of a glass cylinder slipping over the surface. Buffing or rubbing with fine sandpaper is also resorted to for removal of any small imperfections from the grain surface. Chrome leather is strong and durable and for shoe uppers is used almost exclusively. Its preeminence in this field is perhaps largely due to the fact that, once properly dried, the fibers have lost their plasticity, so that on rewetting and again drying out it is not "boardy."

As stated above, where fullness of texture combined with softness is desired, vegetable-tanned leather can be made by employing, not the astringent tannins of heavy leather manufacture, but the milder, less astringent ones. For specialty leathers a number of other tanning agents are available such as alumina, the condensation products of formaldehyde, the oxidation products of certain oils (apparently aldehydic in character), certain new synthetic tanning agents that have been developed in recent years, and the like. However, the backbone of light leather manufacture is still chrome-tanning and, to a distinctly lesser degree, vegetable tanning.

Uses of Leather.

The engineering and industrial uses of leather are small in comparison with its utilization in shoes, gloves, bags, traveling cases, and the like. From the technical point of view its most important characteristic is its combination of strength with remarkable ability to endure repeated flexing. This is best illustrated by its use in belting. Cases are on record in which leather belts have been employed in heavy duty under almost continuous conditions of service for from twenty-five to forty years. However, to secure performance of this sort the belt must be protected from overload. Another case in which leather is employed to the practical exclusion of other materials is the use of soft, vegetable-tanned sheepskin in the construction of dry gas meters. Its superiority for this purpose is based on its ability to withstand indefinite flexing, together with the extraordinary low resistance it offers to the flexing motion. Leather

is unexcelled for use as piston-packing for gasoline and similar liquids. It is important as a fabric for polishing operations and is sometimes employed as a filtering or straining medium. Its use as rawhide was formerly extensive in the construction of small pinions and mechanical parts, but it is here encountering strong competition from vulcanized fiber and similar materials.

In using leather as an engineering material of construction one must remember two important points. In the first place it does not follow Hooke's Law. In the case of belting one usually assumes the elongation proportional to the square root of the load rather than to the first power. This tendency to stretch relatively more at low loads than at high is of course due to its fabric-like structure. In the second place, leather, like all amorphous materials, has a very low elastic limit. Under moderate loads it develops considerable permanent set. However, this effect fades away rapidly under sustained or repeated load. This is the reason for the necessity of relatively frequent tightening of new belts. However, while the tendency to undergo permanent set never completely disappears, it is soon reduced to an almost negligible point. It is on this account that a properly loaded belt, once broken in, will run for months or even years before retightening becomes necessary.

A serious limitation of leather is its sensitiveness to chemical action. It cannot be used in the presence of strong acids or alkalies, or active oxidizing agents. A second limitation is its sensitiveness to moisture. It adsorbs moisture from the air, even at low humidities. Furthermore, this adsorption makes the leather stretch. Thus, vegetable-tanned belting must not be used where it is likely to become thoroughly wet, but the effect on it of mere atmospheric humidity, even though high, is usually unimportant. As pointed out above, chrome leather is far less sensitive to moisture than vegetable-tanned. Both adsorb it and both stretch somewhat in doing so, but chrome leather resists completely the leaching and gelatinizing action which water exerts on the vegetable-tanned product. However, this very capacity of leather to pick up moisture is a major reason for its superiority in the construction of shoes. The inner surface of the upper leather adsorbs the moisture from

the foot. This moisture diffuses through the leather with extraordinary rapidity, almost as though it were dragged along by capillarity. Consequently, the moisture content of the outer surface of the leather is high and the evaporation of the moisture into the air correspondingly rapid. Hence, the moisture given off by the foot as perspiration can escape through a leather shoe almost as rapidly as though the foot were bare.

Various methods of waterproofing leather have been proposed. Most of these involve impregnation with fatty substances. This interferes with flow of liquid water through the voids, but does not greatly retard the capillary movement of adsorbed water along the fibers. In other words, samples of this sort are water-resistant, but in no sense waterproof. Efforts have been made in recent years to impregnate leather with rubber latex as a waterproofing device, but the success of this scheme is not as yet entirely assured. It seems certain that any waterproofing technique which interferes with the capillary movement of the water along the fibers will make the product unfit for upper leather and there is evidence to indicate that the same is true of sole leather, although the foot undoubtedly breathes less through the latter than the former. The ideal treatment would fill the voids with a material which would prevent liquid water from entering and passing through the leather, but would not interfere with the travel of the adsorbed water.

BIBLIOGRAPHY

BENNETT, H. G., *Leather*, Constable and Co., Ltd., London (1926).

LAMB, M. C., *Manufacture of Chrome Leather*, Anglo-American Technical Co., Ltd., London (1923).

MERRY, E. W., *The Chrome Tanning Process*, Harvey, London (1936).

PROCTOR, H. R., *Principles of Leather Manufacture*, Spon, London (1922).

ROGERS, A., *Practical Tanning*, Henry Carey Baird and Co., Inc., New York (1922).

WILSON, J. A., *Chemistry of Leather Manufacture*, Am. Chem. Soc. Monograph, Chem. Cat. Co. (1928–1929).

WILSON, J. A., *Leather*, McGraw-Hill (1931).

WILSON, J. A., *Modern Practice in Leather Manufacture*, Reinhold (1941).

Chapter XVIII

Rubber

Rubber is a material unique in many of its properties. By suitable treatment and manipulation it can be obtained in all states from a hard, non-extensible solid to an extensible material of high elasticity and tensile strength. Its principal and only commercially important source is the milk-like fluid or latex, which exudes from certain trees when they are cut or injured. Chief of these is the "Hevea braziliensis," originally indigenous to Brazil but now grown on plantations situated in countries in the tropical zone, particularly the Dutch East Indies and Malaya. Coagulation of the latex yields the material known as crude rubber, a strong, highly elastic gum, very sensitive to small temperature changes. At low temperatures it becomes hard, but even slightly above room temperature it is soft and tacky, making it of little use for most purposes. It may be plasticized, however, by suitable means, molded, formed, and its plasticity then destroyed by the process known as vulcanization, which results in a product many times stronger and much less sensitive to temperature changes than raw rubber, yet with the other desirable characteristics of the original rubber not greatly affected. In order to understand the changes that take place during this manipulation as a plastic it is essential to gain an insight into the chemical and physical structure of rubber.

Rubber Latex.

The milky fluid known as latex is obtained from the rubber tree by the process known as tapping, involving a small excision from the bark (Fig. 1). A slightly viscous fluid exudes and is collected in a small cup hung on the tree in a suitable position. Flow continues for about two hours after the cut is made and by taking successive pairings of bark at intervals of one or two

days, thus widening the excision and exposing new areas as the old cuts heal over, a supply of latex is obtained day after day.

Rubber latex is an aqueous suspension of spherical or pear-shaped particles of rubber hydrocarbon, about 0.5 to 3.0 μ in diameter, showing rapid Brownian movement. The hydrocarbon constitutes about 35% by weight of normal, freshly tapped latex, but there is present also 2% of natural proteins and small

FIG. 1.   Native Tapping a Rubber Tree in Malaya.*

quantities of resins, sugars, and inorganic salts. The stability of the latex, profoundly influenced by hydrogen-ion concentration, is very great at high pH values, becomes negligible in the range from about pH 6 to pH 3 (the isoelectric point of the particles lies at approximately pH 4.5 to 4.8), and increases again

* Courtesy of Revertex Corporation of America from their pamphlet *Revertex, Concentrated Latex* (1940).

as hydrogen-ion concentration is further lowered. At low hydrogen-ion concentrations (high pH), electrophoresis shows the rubber particles to be negatively charged, since they move to the positive pole in an electric field; on the other side of the isoelectric point, the charge is reversed. The latex suspension is coagulated by electrolytes, the effects paralleling the behavior of suspensoids. The precipitating action of salts on fresh latex with its pH of about 7 depends mainly on the valence of the cation. Thus, the coagulating effect on a latex diluted to a rubber concentration of 1% or below of 0.0006 equivalents per liter of $Al_2(SO_4)_3$ requires 40 times the equivalent concentration of $MgSO_4$ and about 1000 times the concentration of alkali chloride. These differences confirm the nature of the charge on the particles (Hardy-Schulze rule, p. 125). However, while concentrated latex requires for coagulation far more aluminum sulfate than dilute, little or no more of the common divalent cations is needed.* Hydrogen ion is unique in its action; thus, the equivalent concentration of acetic acid necessary for coagulation of dilute latex is but little more than that of $Al_2(SO_4)_3$ but a 50-fold increase in rubber concentration requires a 10-fold increase in acetic acid concentration. The latex has a lower pH range of maximum stability, extending from approximately pH 3.0 to 1.2. High concentrations of strong mineral acids such as hydrochloric induce increased coagulation, presumably due to their effects as concentrated electrolytes. Dehydrating agents are also strong coagulants, 50 volume per cent of ethyl alcohol causing rapid flocculation.

In the production of crude rubber on the plantation the latex is usually coagulated by the addition of formic or acetic acid,† and the resulting spongy mass, in which the serum is

---

* Thus, 50-fold increase in rubber concentration requires 25-fold $Al_2(SO_4)_3$ but only 1.3-fold the $MgSO_4$. However, copper sulfate is out of line with the other divalent cations, closely paralleling aluminum sulfate. Probably a large fraction of the aluminum and cupric ions, rather than merely modifying the character of the external solution, combines directly with the protein.

† Freshly tapped latex has a pH of about 7.0 but, on keeping, rapid action by bacteria sets in, the pH drops suddenly to about 6.2 or 6.5 and simultaneously coagulation occurs, the whole process taking but a few hours. Titration by alkali shows that a considerable amount of free acid is formed during this change, no noticeable effect being produced on the pH of the latex in the early stages because of the natural buffers present in the sol. Finally, however, the acid makes itself felt and the protein is taken into the range of instability, with consequent coagulation.

enmeshed, is passed between rollers. In this way most of the serum is squeezed out and a sheet is formed, from which, after washing, the remaining water can be removed by drying. Smoked sheet is prepared by passing the still wet coagulum through a series of smooth, even-speed rolls, followed by one pass through embossing rolls, drying being conducted at 40 to 50° C. in an atmosphere of smoke produced from such materials as green wood, cocoanut shells, etc. Due to phenolic derivatives in the smoke, this treatment not only helps to preserve the sheets against action by bacteria but gives them a uniform brown color. In the production of crêpe rubber the coagulum is torn apart under a stream of water by rolls moving at unequal speeds and is then pressed into a sheet, which is air-dried at room temperature.*

Manipulation of Latex.

During the last quarter of a century it has become customary to use latex directly for the fabrication of certain types of rubber goods and articles (p. 437). Since little manufacturing is done in the Far East, it is necessary to stabilize the latex against the development of bacterial coagulation already mentioned and, in order to avoid paying for transportation of excessive quantities of water, to concentrate it. For stabilization, many alkaline materials, e.g., borax, sodium and potassium hydroxides, can be added to the latex, thus preventing any increase in hydrogen-ion concentration with accompanying subsequent coagulation. However, ammonia is the most successful and widely used addition agent for this purpose, about 0.5 to 1.0% by weight being added.

Commercial concentration of rubber latex is usually effected by one of three processes. By treatment in a centrifuge of the cream separator type at 5000 R.P.M. or upwards, ammonia latex may be separated into a concentrate containing up to about 65% rubber and a skim containing about 10 to 15% solids. Such centrifuged latex differs from ordinary latex in that it has been partially purified, much of the protein and resinous content of the original latex remaining in the skim. This, as will become clear later (p. 430), affects its vulcanization properties and increases the water resistance of the vulcanized product. Concentrated latex may also be obtained by

* It is sometimes used without further treatment in shoe soles.

"creaming," or Traube's process. In this, a dilute solution (0.1 to 0.5%) of a highly hydrated colloid, such as Irish moss, gum tragacanth or tragon seed gum, is added, whereupon the solution separates into two layers, the upper containing 50 to 60% of rubber. Under the microscope the creaming agent can be seen to stop the Brownian movement of the rubber particles, which clump together and form agglomerates. On dilution of the cream, reversal of this process occurs and Brownian motion again develops. The enlarged agglomerated particles, of specific gravity less than 1, readily rise to the surface of the latex. Like centrifuged latex, creamed latex has been partly purified from the proteins and resins originally present, but still contains the creaming agent, which may influence the latex in subsequent manufacturing operations. In the third process, a concentrated latex containing about 70% rubber is obtained by evaporation. In order to prevent coagulation, a protective colloid, usually soap together with additional alkali, is added to the latex, which is then concentrated in evaporators designed to avoid skin formation. Evaporated latex naturally contains all the constituents of the original latex.

Chemical Properties of Rubber.

Analyses of various types of high-grade crude rubber show that it contains a fraction of a per cent of ash, from 2 to 3.5% of protein, and from 1.5 to 3.5% of resins (soluble in acetone). The more completely the above impurities are removed, the more closely does the remainder approach the composition represented by the empirical formula, C_5H_8. Osmotic pressure and viscosity measurements of solutions of the rubber hydrocarbon, using the method of Staudinger (p. 165), indicate that the molecular weight is very large, of the order of 100,000 to 350,000, depending to some extent upon the previous treatment of the rubber, a result confirmed by ultracentrifuge measurements.

Destructive distillation throws light upon the chemical structure of rubber. When heated in air, it softens at 120° C. and changes to a thick brown oil, which never regains the properties of rubber on cooling. At higher temperatures, approximately 300° C., cracking occurs, and upwards of 90% of volatile hydrocarbons, based on the weight of original rubber, have been obtained. These hydrocarbons consist of small amounts of isoprene,* with large quantities of isoprene polymers. This, together with the fact that rubber hydrocarbon contains one double bond for each isoprene unit, makes it probable that

* By very rapid distillation in vacuo up to 23% of isoprene can be obtained.

rubber is a chain polymer of isoprene of the nature of the compounds discussed on p. 140. In other words, rubber has the chemical formula $(C_5H_8)_n$, where n is large and variable, probably differing from molecule to molecule even in a single sample of rubber (cf. cellulose, p. 152 and see also p. 154).

$$CH_2\!\!=\!\!C\!\!-\!\!C\!\!=\!\!CH_2 \;\rightarrow\; -CH_2\!\!-\!\!C\!\!=\!\!CH\!\!-\!\!CH_2\!\!-\!\!CH_2\!\!-\!\!C\!\!=\!\!CH\!\!-\!\!CH_2\!\!-$$

with H above the second carbon and CH_3 below on the left (isoprene), and CH_3 above on the left center and CH_3 below on the right (rubber).
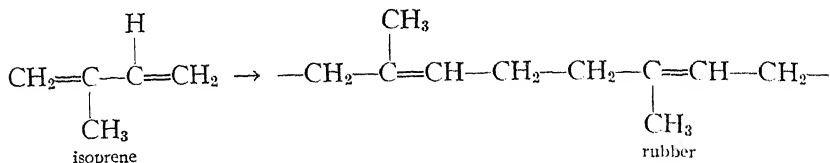
This structure is supported by a number of confirmatory facts. Rubber combines with both bromine and iodine to give compounds $(C_5H_8 X_2)_n$, corresponding to one double bond per C_5H_8 unit. With ozone, rubber gives an ozonide, which on decomposition with water gives laevulinic aldehyde and its oxidation product laevulinic acid, over 90% of the original rubber being accountable for in these forms. This is readily represented by the above formula for rubber:

$$-CH_2\!\!-\!\!C\!\!=\!\!CH\!\!-\!\!CH_2\!\!-\!\!CH_2\!\!-\!\!C\!\!=\!\!CH\!\!-\!\!CH_2\!\!-$$

with CH_3 above each double-bonded carbon.

$$\downarrow \quad O_3 \text{ followed by } H_2O$$

$$\underset{O}{\overset{H}{\diagdown}}C\!\!-\!\!CH_2\!\!-\!\!CH_2\!\!-\!\!C\underset{O}{\overset{CH_3}{\diagup}}$$

laevulinic aldehyde
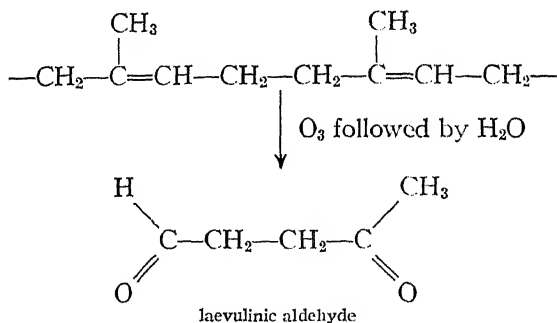
Furthermore, isoprene has a marked tendency to polymerization, and, under favorable conditions, will polymerize to a product resembling rubber although not identical with it.* It seems safe to infer that rubber hydrocarbon is a polymer of isoprene of the chain type.

---

* It may be well to state at this point that no substance completely resembling rubber in all its properties, physical and chemical, has ever been prepared. In other words, the term "synthetic rubber" is applied to any product having the extraordinary extensibility with high tensile strength, which is characteristic of rubber, particularly if it can also be vulcanized.

Physical Properties of Raw Rubber.

Raw rubber differs greatly from most materials in its be-havior on deformation.   If a sample be placed under tension in a testing machine capable of measuring simultaneously both load and elongation, the result-ing stress-strain diagram (Fig. 2) is concave to the stress axis, *i.e.*, the rubber becomes increasingly difficult to elongate as the stress increases (Fig. 2).   This offers a sharp contrast to the behav-ior of a material such as steel (Fig. 2, p. 277, noting that the scale is logarithmic).   Further-more, the elongation before rup-ture is very great, up to 800 to 1000% of the unstretched length.   On removal of load, the stress-strain curve lies below that corresponding to elongation, *i.e.*, there is a hysteresis loss of me-chanical energy in the cycle. Even more significant is the effect of time on the elongation.

FIG. 2.   Stress-Strain Diagrams for Rubber.   (The temperature on each curve is that of test.   The in-dividual points are breaking loads.)

Thus, if the stress be held constant at a load below the breaking strength, the sample will elongate slowly for an indefinite length of time and, when the load is great enough, will ultimately rupture.   If the load be taken off before rupture, the sample never contracts to its original length; there is an immediate contraction equivalent to a high percentage of the initial elonga-tion, followed by a slow further contraction, which, however, is usually only a small fraction of that initially occurring.   The remaining or permanent set of raw rubber can amount to as much as 30% of the total elongation under load.   If a sample thus treated be loaded again, the phenomena repeat themselves, an additional permanent set results and, if the process be re-peated a sufficient number of times, the sample finally breaks.

On stretching a sample of raw rubber its physical properties change in four significant respects. In the first place it evolves heat, a phenomenon known as the Gough-Joule or Joule effect. Furthermore, this energy change is to a considerable degree reversible, the sample dropping in temperature when allowed to contract. In the second place, the stretched sample shows an increase in density from about 0.937 at zero elongation to 0.950 at 1000% elongation. Third, the stretched rubber, initially nearly transparent, becomes white and opalescent in appearance. Finally, whereas in the unstretched condition raw rubber gives an X-ray diffraction pattern typical of liquids and most amorphous solids (Fig. 3), under an elongation as low as 80% a definite fiber diagram begins to make its appearance, the spots being situated on the rings of the frozen rubber pattern described below. The positions of these spots remain unchanged during increasing elongation, although their sharpness increases. The critical elongation, 80%, at which the X-ray fiber pattern first makes its appearance is also the same as that at which the Joule effect begins to develop, little or no heat being evolved below this elongation.

Raw rubber is unique in its response to temperature changes, the observed effects being most illuminating in coming to a decision as to its structure. As already stated, it becomes sticky and tacky on heating, even under conditions such that any marked chemical change seems unlikely. On cooling, it hardens, and at very low temperatures becomes quite brittle. In other words, rubber is thermoplastic, with a high temperature coefficient of plasticity. If held for considerable time at moderately low temperatures (10° C. and somewhat below), it becomes progressively stiffer, harder, and opaque, ultimately developing an X-ray ring diffraction pattern, such as is characteristic of incompletely oriented crystallites.

When a sample of raw rubber is stretched at ordinary temperatures, chilled, and the load then removed, recovery is greatly retarded and less complete. This effect develops at temperatures as high as 10° C. and is very marked at 0° C. On the other hand, if the stretched, chilled sample be allowed to warm up, it immediately recovers its original length with no

more permanent set than would occur under ordinary conditions. In contrast, when a sample of rubber, previously stretched at normal temperatures, is heated somewhat, as by plunging in hot water, it loses in large degree its capacity for recovery, but if cooled, it can by repetition of the process of alternate extension and heating be stretched further, elongations of 100-fold the original length being achieved without trouble. Rubber thus treated is said to be racked, a term sometimes also applied to stretched, chilled rubber. Racked rubber shows a spot-type X-ray diffraction pattern, characteristic of oriented and fibrous materials. Finally, racked rubber chilled to very low temperatures becomes brittle and on impact breaks up into fibrous masses not unlike asbestos, while unstretched rubber under the same conditions pulverizes like glass. The structure of rubber is greatly influenced by its thermal and mechanical history.

Granting that rubber consists of long molecular chains of polymerized isoprene, its X-ray diagram (Fig. 3) indicates that

(a)   (b)
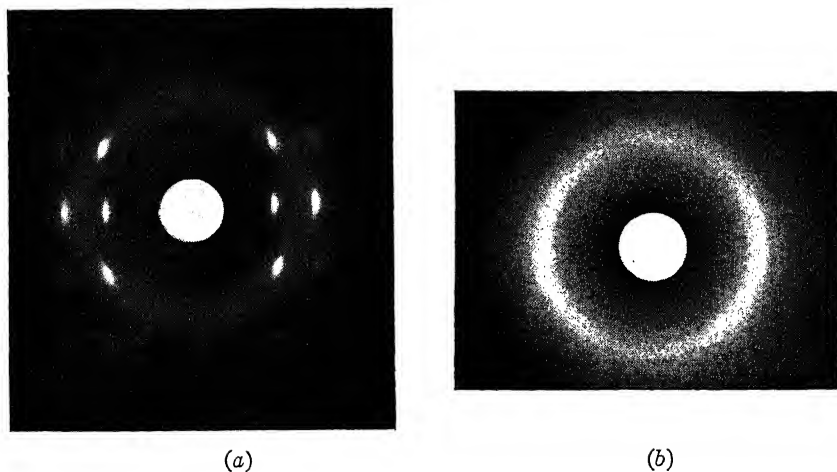
FIG. 3.   X-Ray Diagrams of (a) Stretched and (b) Unstretched Rubber.*

they are practically unoriented. The unvulcanized rubber must be conceived as consisting of snarls of molecular chains or threads, but snarls in which the individual threads are closely

* Courtesy of Dr. B. E. Warren, Department of Physics, Massachusetts Institute of Technology.

packed and intertwined.* However, these molecular chains are unique in that the successive elements (*i.e.*, the carbon atoms in the chain) are attached at tetrahedral angles, but with a considerable degree of freedom of rotation of the chain around each such axis of attachment, subject to substantial retention of the angle of divergence.† Study of the situation will make it clear that a mass of chains of this type, irregularly packed, can undergo very great distortion without rupture, provided one can assume flexing and rearrangement of the individual molecular chains, largely by rotation at the points of linkage of the carbon atoms, together with a certain freedom of slippage of the molecules over each other at all points of contact. This slippage will allow readjustment of the chain arrangement on elongation and yet offer a high degree of resistance to rupture when the molecules have pulled out into positions more or less parallel, because adjacent molecules now have a large area or many points of contact, despite the fact that the resistance to slip (*i.e.*, intermolecular friction) may be small at each specific contact point. The arrangement must approximate that indicated in Fig. 4, the axes of some threads being more or less parallel to the plane
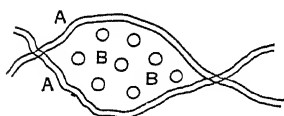
FIG. 4.

of the diagram, while other threads, passing between the loops of the former, have their axes perpendicular.

When a sample of raw rubber is submitted to a moderate elongation, the loops *AA* of Fig. 4 elongate in the direction of pull, reducing the cross section of the loops and distorting the molecular chains *BB*, perhaps both compressing and rolling them slightly. There will be a limited amount of slip of the molecules past each other, but both the molecular distortions and the compressive action on the chains *B* will exert a restoring force, opposing the elongation and tending to bring the sample back to its original shape. A part of the distortions of the molecular chains may even be due to actual distortion of the

* The molecules cannot be long rods, since, if unoriented, this would involve relatively large voids between them. This, in turn, would mean low density, whereas the density of rubber is high, relative to what one would expect for close packing of molecules of its type. The only alternative is to assume the molecular chains wrapped around each other in more or less sinusoidal shapes.

† See *J. Am. Chem. Soc.* **60**, 494–5 (1938).

valence bonds of the carbon atoms, owing to which the chains would be expected to have a zigzag or conceivably even helical shape. As elongation of the sample increases, all the molecular chains tend to be pulled out into relatively parallel positions. The chains B caught in the loops of A would still be held, but their ends outside the loop would tend to come into parallelism with A, all in the direction of pull.

A picture of this sort gives a satisfactory explanation of the more important physical properties of raw rubber. The intermolecular friction is high but insufficiently high to prevent considerable permanent slip between adjacent molecules when under sufficient strain. This slip is equivalent to flow of the rubber as a liquid, and once it has occurred there is no significant force tending to restore the original configuration. In other words, one would expect rubber to exhibit permanent set, increasing rapidly with the magnitude of the distorting forces and their time of application. At high elongation, the parallelism of all the molecular chains makes further elongation, due to closing of loops such as those between the molecules AA of Fig. 4, impossible. This greatly increases resistance to further movement, offering an explanation of the marked increase of stiffness with elongation. The parallelism of the molecular chains when stretched accounts for the X-ray diagrams. This parallelism is insufficient to give observable X-ray diffraction effects until considerable elongation has occurred, but the increased parallelism of orientation beyond this point (about 80%) results in the marked sharpening of the diffraction diagrams found. As the molecular chains are brought into parallel positions, the opportunity for effective action of the van der Waals intermolecular attractive forces increases, drawing the molecules closer together and increasing the density. The operation of these forces liberates energy, which exhibits itself as the Joule effect.

The only way in which the rubber can elongate to any considerable degree involves slip of the long molecules over and past each other. However, because at all points of contact the molecules are attracted together by van der Waals' forces, there must be intermolecular friction requiring an externally imposed

force to overcome it. All work done by such forces is dissipated by conversion into heat and is therefore non-recoverable as mechanical energy. This lost work constitutes the hysteresis effect always observed in the distortion of rubber.

As the temperature of a sample of rubber rises, the rubber expands. In consequence, the molecular interattractions are greatly reduced (p. 3). This in turn reduces the friction between chains at all points of contact, and hence also reduces the energy lost by the hysteresis (p. 424) accompanying the distortion cycle. The effect parallels the decrease in viscosity of a liquid with rise in temperature (p. 29 *et seq.*).

If stretched rubber be heated, the reduction in interparticle friction can allow the molecules to slip sufficiently to adjust themselves to the stretched condition. This eliminates the strains, which alone could cause recovery of shape on release of load. Repetition of such elongation followed by heat treatment ultimately results not only in very great elongation of the sample but also in almost perfect, parallel orientation of the molecular chains as evidenced by the X-ray diagram.

While release of the load on an ordinarily stretched sample involves recovery of shape, due to the resilient effect of the compressive and other stresses described above, this can occur only because of the necessary though limited slip between the rubber molecules. If now the mass be chilled, particularly while in a stretched condition, thermal agitation of the chains decreases and the net intermolecular friction increases, tending to freeze the mass permanently into its stretched shape, despite the internal strains attempting to restore it. When stretched rubber is chilled to very low temperatures and disintegrated by impact, its fundamentally fibrous structure discloses itself in the fibrous character of the fragments. Finally, if unstretched rubber be held at a proper low temperature level there would appear to be actual crystallization, or at least, what is equivalent to it, a localized orientation of the molecular chains. This shows up in the evidences of non-homogeneity, in the X-ray diagram, and in the nature of the fragments obtained by fragmentation at very low temperatures.*

---

* For discussions of the molecular structure of vulcanized and loaded rubbers, see pp. 417–427 and 430–437.

Milling.

Crude rubber is insufficiently plastic for satisfactory mechanical manipulation in the processes of manufacture. Although it can be plasticized by heat or solvents, the former, probably due to chemical action, results in deterioration, while the resistance of the raw rubber to solution and viscosity of the resultant liquid preclude the effective use of the latter. However, it is found that if rubber be compelled to flow under high shearing stress, its plasticity can be increased to any desired extent without serious loss in quality of the ultimate product. This operation is known as milling or mastication.

Commercially, milling is performed on a roll mill or in a Banbury mixer. The former (Fig. 5) consists of two parallel, horizontal rolls of chilled cast iron, usually 60 or 80 ins. long and
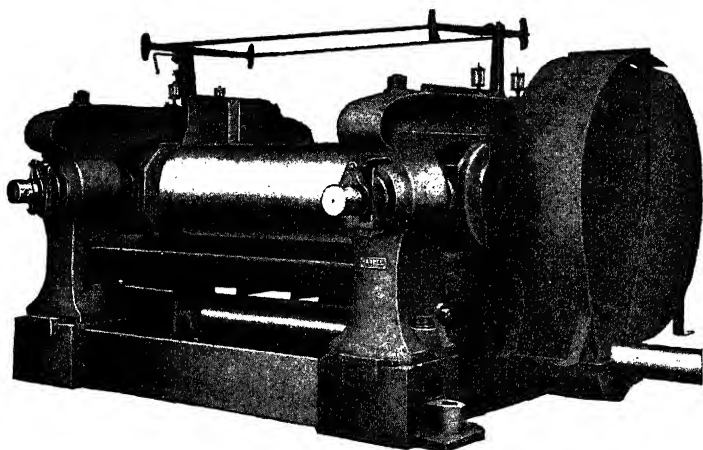
FIG. 5. Roll Mill.*

20 to 25 ins. in diameter. The back roll is driven, *e.g.*, by an electric motor, and is coupled to the front roll through a set of gears, so that it runs faster than the latter. Milling is a batch operation, sufficient stock being kept above the rolls to maintain a considerable "bank" of material in the nip between them. The differential surface speed, roughly 25%, facilitates mixing by keeping this bank rolling. As milling starts, ordinary

* *Lubrication of Farrel-Birmingham Machinery*, Bull. No. 52, p. 9, Farrel-Birmingham Co., Ansonia, Conn.

raw rubber will not stick to the rolls as it comes through but must be picked up by hand and returned to the top, but at a later stage the sheet issuing from the point of nip adheres to the front roll and returns automatically to the bank. Because the bank tends to thin out from the middle to the edges, the flow through the latter and consequently the plasticization also are less. To get uniformity of mastication, the sheet returning on the front roll is, from time to time, cut off by the operator, rolled up, and piled into the middle of the bank. The rolls are run at the maximum allowable speed, dictated by the tolerance of the stock to temperature rise. The whole operation takes from thirty to sixty minutes for each batch, depending upon the type of stock being used.

There is sufficient adhesion of the rubber to the rolls at the point of nip to have little or no slip at the roll surface. This
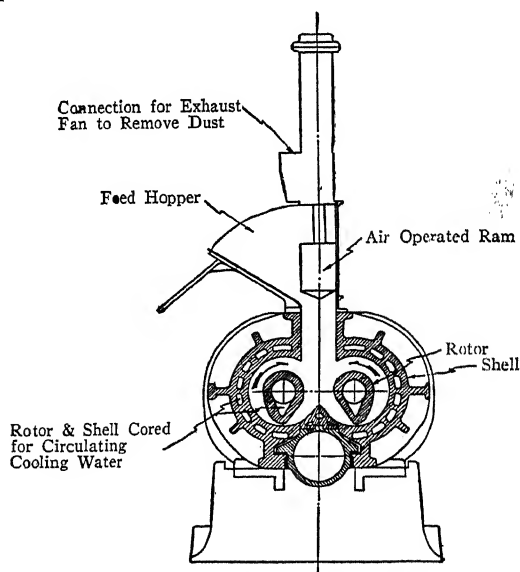
FIG. 6. Banbury Mixer.*

requires the imposition of a shearing force, the magnitude of which depends on the plasticity of the stock, and which is controlled by the speed of the rolls and the distance between them. In the case of raw rubber, this is extremely high. It is found

* *Banbury Mixers*, Bull. No. 171. Fig. 7, Farrel-Birmingham Co., Ansonia, Conn.

that rubber submitted to these conditions undergoes a change, resulting in rapid increase in plasticity. This in turn reduces the force necessary to maintain the rate of flow between the rolls. As the shearing force drops, the rate of increase in plasticity of the stock also drops, and in time the rubber reaches a condition in which the force necessary to maintain its rate of flow through the mill is insufficient to cause further plasticization, i.e., it reaches an equilibrium state. This results in a several-fold variation in power load, the evils of which can be minimized by driving two or more mills from a single motor. However, the power consumption is always large.

In the Banbury mixer (Fig. 6) two kneading arms or rotors revolve inside a mixing chamber, the wall of which is concentric to the rotors. In this way a thin wedge of rubber is continuously formed and submitted to a shearing action, while the operation of the rotors at slightly different speeds insures good mixing of the batch.

Not only is masticated rubber more plastic, but it swells in solvents more rapidly, disperses in them more readily, and the viscosity of its solutions is greatly reduced (Fig. 7), all these effects increasing progressively with the severity of milling. Furthermore, the degree of plasticization is greater at lower temperatures of milling, corresponding to the greater mechanical work expended upon the stiffer stock (Fig. 8).* For this reason, both roll mills and Banbury mixers are water-cooled and temperature control of the stock on the mill is one of the important methods of quality regulation. This transformation of the stock on milling is known in the industry as "breakdown."
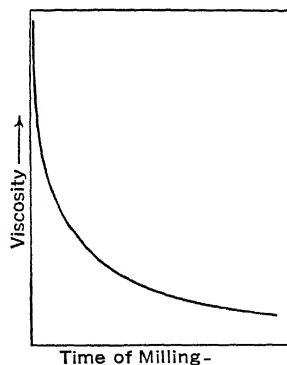
Fig. 7. Change in Viscosity of Rubber Solutions on Milling.

* The arbitrary measure of plasticity used in the experimental determinations of Figs. 8 and 9 is the set of a cylindrical sample submitted to a constant axial compressive load for a constant time, after which the load is removed and recovery allowed to occur for an equal time, expressed as a percentage of the average length of the loaded and unloaded sample. This figure is conventionally divided by 2. The measurements were made at 100° C.

From the general characteristics of high molecular weight materials (p. 142), the obvious inference is that breakdown is caused by the shortening of molecular chain length of the hydrocarbon. Indeed, this explanation has long been held, although many chemists hesitated to accept it because of the improbability of rupture of the primary bonds of the carbon linkages by relatively weak mechanical forces.

Developments of the last decade have offered a clue to the solution of the problem. It was discovered by Cotton that the mechanical manipulation of raw rubber has little or no effect upon its plasticity or other physical properties if conducted in absence of oxygen. Figure 9 shows typical results obtained when milling crêpe rubber in various gases, the mill being enclosed so that its atmosphere could be controlled at will. During mastication the rubber absorbs small but definite amounts of oxygen and considerable quantities of peroxides are formed, presumably at the double bonds. There can be no doubt that oxygen attacks the rubber hydrocarbon chain under suitable conditions and that mechanical stress favors the attack. The reason for the attack occurring only under conditions of manipulation is far from clear (cf. wool, p. 500). Perhaps the double bonds are more reactive chemically when under mechanical strain, or the oxygen attack may be due to electrical charge induced by the friction of milling, e.g., by ozone formation. In any event, the hydrocarbon chains apparently are ruptured, the initial step in the process being oxidation of one or more of the double bonds by some reaction speeded up by mechanical strain on the material. This point of view is confirmed by the fact that chain hydrocarbons, such as polymerized isobutylene, which

FIG. 8.   Effect of Temperature on Milling.*

* SYSKO, A. J., *Mass, Inst, Tech. Thesis* (1933).

contain no double bonds distributed along the chain, show negligible breakdown on milling and a resistance to oxygen attack which is far higher than that of rubber.
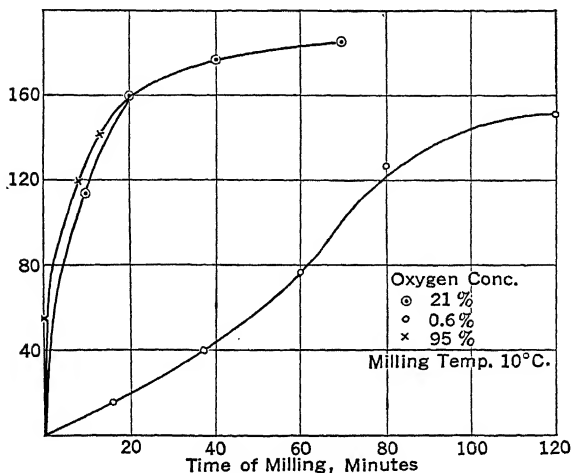
FIG. 9.   Effect of Oxygen Concentration on Milling.*

In the early stages of milling, it is common practice to introduce into rubber various materials, known as softeners, such as stearic acid, pine oil, and the like.   These increase the plasticity of the rubber, reducing the time and power required for ultimate breakdown, and are also said to have a favorable effect upon the dispersion of the fillers, *e.g.*, carbon black, subsequently compounded with the rubber.   It seems likely that they act predominantly as internal lubricants, due to mutual solubility with the rubber.†   Their major value probably lies in ability to secure the desired degree of plasticity on the mill with less breakdown of the rubber itself, *i.e.*, less reduction in chain length than would otherwise be necessary.‡

Both roll mills and Banbury mixers, after completing the breakdown of the raw rubber, are used as mixing mills for incorporating into the rubber other components of the stock.

* SYSKO, A. J., *Mass. Inst. Tech. Thesis* (1933).
   † Thus molten stearic acid dissolves a considerable quantity of rubber, which can be subsequently precipitated by alcohol.
   ‡ Other softeners differ in type and presumably in mechanism.   Small quantities of phenylhydrazine are extraordinarily effective in softening rubber and reducing milling time.

These may be summarized under the heads: vulcanizing agents (almost always sulfur), reinforcing agents, fillers, anti-oxidants, and accelerators. (The purposes of these will be made clear in the following pages.) After homogeneity is attained, the rubber is ready to be shaped before heating for vulcanization.

Shaping and Molding.

After compounding or mixing in the various ingredients, a plastic, tacky mass is obtained which can be manipulated in various ways according to the type of goods being manufactured. For many articles rubber sheets of uniform thickness are first produced by calendering, in which the plastic mass of rubber is fed between the upper two of three rollers, situated vertically one above the other. The middle roll, revolving in the opposite direction to the other two, carries the rubber forward and delivers the sheet between itself and the lower roll in uniform thickness. The plastic rubber mix may also be extruded through a die when continuous lengths of regular shape are desired. Where molds are used, the rubber mix is placed between the two halves of the mold, which is then closed in a press, usually hydraulic, and heated to carry out the subsequent operation of vulcanization. For small articles, the plates of the press itself are heated, either electrically or by steam, the molds being placed between them. More even heating is secured by introducing the steam around the mold itself; this requires the use of an autoclave press, which is particularly necessary for large objects. The rubber takes a sharp impression of the mold and a well-defined vulcanized article is produced. Hollow goods can be made by lining the mold with calendered sheet, a small quantity of ammonium carbonate or nitrite being placed inside before closing; on heating, the salt dissociates, producing sufficient internal pressure to give a clear cut article. Some products, such as rubber tubing, can be supported on metal forms or mandrels, the outer surface being wrapped in cloth, or, alternatively, the whole article is embedded in an inert powder, such as French chalk, and submitted to the action of live steam in autoclaves, the process being known as "open steam" vulcanization. "Dry heat" vulcanization consists in heating the formed

articles in warm air, and, because the cured rubber is apt to be porous, is used only for thin goods, *e.g.*, footwear, where tensile strength is of minor importance. Vulcanization in hot water is also sometimes employed.

Vulcanization.

The usefulness of raw rubber is limited because of its high sensitivity to temperature changes, large permanent set, and relatively low strength. Milled rubber is even weaker than raw rubber and is almost completely useless as such. In 1839 Goodyear found that its properties are greatly improved by milling small quantities of sulfur into masticated rubber and heating the mix at 130 to 150° C. for several hours. The mechanical properties of the rubber are transformed, the tensile strength at break being increased about seven-fold (Fig. 2); the rubber is far less thermoplastic, so that it can be used successfully over a much wider range of temperature than can raw rubber; the solubility or swelling of the rubber in organic solvents is decreased; and, finally, the permanent set of the rubber is greatly reduced. The change is accompanied, as shown below, by evidence of chemical combination of the sulfur with the rubber. Two words, "cure" and "vulcanization," are used to describe it. Although they are frequently treated as synonyms, it seems desirable, as is often done, to reserve the word "cure" to describe the process of the transformation in physical properties (as in *overcure, undercure,* etc.), employing the term "vulcanization" for the chemical changes involved.* Both words are used to refer to the process as a plant operation.

Extensive work has been carried out on the vulcanization reaction with sulfur without leading to a definite conclusion as to the precise mechanism involved. Certain facts are, however, well substantiated. After vulcanization, extraction with acetone or other solvents fails to remove all the sulfur, as does even more drastic chemical treatment, such as heating with metallic copper. In other words, the change is in large degree irreversible. The fact that vulcanized rubber cannot in general be

---

* This phraseology is widely, although not universally, employed. The one term which is never used equivocally is "coefficient of vulcanization" (p. 418).

used to vulcanize fresh rubber also indicates that the change is not merely catalytic. Evolution of hydrogen sulfide is usually small; hence the reaction must be predominantly sulfur addition. Unsaturation as measured by halogen addition always decreases, usually involving disappearance of one double bond for each sulfur atom combined, although recent data indicate that the reduction is sometimes less. This eliminates simple hydrocarbon polymerization of the type:

$$
\begin{array}{cc}
\underset{\substack{|\\ \text{CH}_3}}{\text{CH}_3} \;\; \underset{\substack{|\\ \text{H}}}{\text{H}} & \underset{\substack{|\\ \text{CH}_3}}{\text{CH}_3} \;\; \underset{\substack{|\\ \text{H}}}{\text{H}} \\
\end{array}
$$

—CH₂—C = C—CH₂—   —CH₂—C — C—CH₂—
          +   →
—CH₂—C = C—CH₂—   —CH₂—C — C—CH₂—
          | | | |
         CH₃ H   CH₃ H
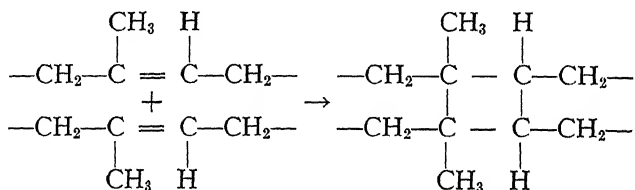
The amount of combined sulfur, expressed in grams of sulfur per 100 gms. of raw rubber, the so-called vulcanization coefficient, rarely if ever exceeds 47, corresponding to one sulfur atom for each double bond in the C_5H_8 unit, and is usually far less. Furthermore, the temperature coefficient of the change, as measured by the ratio of the rates of vulcanization at temperatures 10° C. apart, is about 2.5, of the same order as that of typical chemical reactions. The change is thus unmistakably chemical, probably taking place by sulfur addition at the double bonds. Perhaps the simplest mechanism compatible with these chemical facts is the assumption that a sulfur atom adds directly to each double bond. However, considering the influence of chemical structure on the behavior of high molecular weight materials (p. 155), it seems impossible to explain on so simple a hypothesis the profound changes in physical properties of rubber induced by almost inconsequential additions of sulfur. To understand vulcanization it therefore becomes necessary to study the corresponding physical changes during cure.

The most striking point about the vulcanization of rubber is that, under suitable conditions, chemical addition of less than 0.5% of sulfur to the rubber hydrocarbon (corresponding to less than 1% of the sulfur theoretically necessary to saturate its double bonds) can be sufficient to effect a complete cure, increasing the tensile strength to many times that of the milled or even

the original raw rubber. Plasticity is largely eliminated and the effect of temperature on the product is greatly reduced. Perhaps equally significant is the fact that the tensile strength of raw unmilled rubber can be increased to almost the same degree by merely lowering its temperature to approximately $-50°$ C. Finally, vulcanization very greatly reduces the swelling of the rubber in solvents, as illustrated by Table I.

TABLE I.   SWELLING OF VARIOUS VULCANIZATES
IN BENZENE (APPROXIMATE)

Vulcanization Coefficient	Swelling after 24 hrs. (Final Volume of 1 c.cm. of Original Rubber)
0.	30
1.2	13
6.0	6
39.0	1.4
47.0	1.1

All these facts suggest that the chemical combination of sulfur with rubber is no mere addition of a few sulfur atoms along the hydrocarbon chain, which should result in but little change in structure, but that it involves an extraordinary increase in complexity of structure, only possible by cross-bridging between adjacent chains. The strongest argument for this interpretation perhaps is found, not in the complicated behavior of rubber itself, but in the simpler case of the polymerization of styrene (p. 157). It will be remembered that pure styrene cannot be polymerized to an insoluble product. However, the introduction of as little as 0.1% of an impurity (divinyl benzene) chemically capable of inducing cross-linkages between the styrene polymer chains makes it possible to polymerize it to a product which swells far less in solvents and will not disperse, i.e., a product behaving in this respect exactly like vulcanized rubber. The increase in tensile strength of rubber as temperature is lowered seems almost certainly due to reduction in slippage past each other of the rubber molecular chains. A relatively small amount of chemical cross-bridging would be expected likewise to minimize such slippage.

The mechanism of sulfur cross-linking is still unknown. However, keeping in mind the stoichiometric relation between reduction in double bonds and sulfur added, it would appear that cross-bridging can be explained most readily on the assumption of some structure * similar to one of the following possibilities:

```
   CH₃  H                    CH₃    H
    |   |                     |     |
   —C — C—                  —C  —  C—
    |   |                      \   /
    S   S                       S
    |   |                       ‖
   —C — C—                      S
    |   |                      /   \
   CH₃  H                   —C  —  C—
                              |     |
                             CH₃    H
```
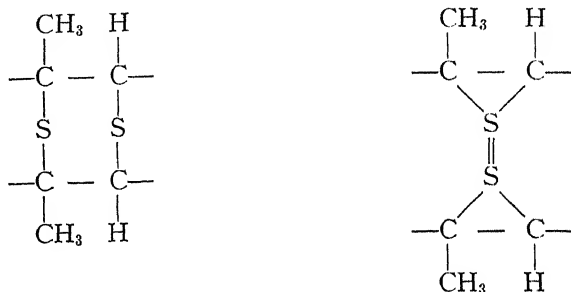
Figure 10 plots tensile strength and elongation at break, for a mix containing 100 parts rubber and 50 parts sulfur cured for varying times at 142° C., against coefficient of vulcanization. Up to a coefficient of nearly 12, the rubber has a moderately high tensile strength and is readily extensible; from approximately 15 to 25, a leathery intermediate product, possessing little tensile strength, is obtained; finally, with high combined sulfur, a hard rubber with high tensile strength but little extensibility is formed. At the same time the solubility and swelling of the rubber decrease with increasing combined sulfur, as shown by Table I. As a consequence of this behavior, commercial rubber lies either in the zone of low sulfur content, so-called soft rubber, or contains 25 to 35% combined sulfur, as ebonite or hard rubber.

The differing effects of liquids on the tensile strength of cellulose in the form of threads and of paper were discussed on p. 347. That soaking with water raises the tensile strength of the former while lowering that of the latter was explained as due to the fact that water makes the fiber slippery. In the thread the fibers are long enough and sufficiently intertwined so that, while slippage induced by water allows all of the threads to come under substantially maximum load, the twist prevents ultimate

---

* It must be pointed out that analogous cases where undoubted sulfur cross-links occur between simple organic compounds are extremely rare.

rupture by slip.   In paper, on the other hand, water makes the
fibers practically fall apart.   Unvulcanized rubber is in some
ways intermediate in its behavior between thread and paper.
While snarling of the molec-
ular chains is sufficient to
retard ultimate slip, slip to
the point of rupture will
occur under sufficient load.
Addition of sulfur to the
rubber is equivalent to fas-
tening the snarls together at
certain of their points of
contact by the cross-bridg-
ing between the molecular
chains.   This strictly limits
the amount of molecular slip
which is possible without
molecular rupture.   Rela-
tively little of this increased
resistance to slip is necessary
greatly to increase the ten-

Fig. 10.   Variation in Properties with
Coefficient of Vulcanization.*

sile strength of the rubber without much change in its elongation
under load.   If, however, sulfur addition with its consequent
cross-bridging be carried too far, it can interfere with adjustment
of the load by slip between the molecular chains, just as removal
of the water from the cellulose thread reduces the slip of the
fibers and in consequence lessens the tensile strength.†   This
is presumably the cause of the low tensiles at intermediate co-
efficients of vulcanization found under the specific conditions of
cure of Fig. 10.   If vulcanization be continued sufficiently, the
whole mass of rubber is interlocked into a rigid structure, in
which any adjustment by slip under load is impossible, but in
which the mass is held together by strong primary valence bonds,
giving it considerable mechanical strength.   Nevertheless, be-
cause of the random orientation of the linkages and the impossi-

* SMITH, C. C., *Mass. Inst. Tech. Thesis* (1932).
† This analogy must not be pressed too far.   Thus, in the twine the fibers are initially
nearly parallel so that elongation under load is slight, in sharp contrast to the snarled
condition of the hydrocarbon molecules of unstretched rubber.

bility of adjustment of load on the various molecular bonds under strain, the tensile strength of this highly vulcanized product, figured on actual cross section at break, is only a fraction (roughly 40%) of the strength of soft rubber, in which, at break, the molecular chains are combed out into parallel positions.

The influence of temperature on the elastic properties of rubber throws light on the structure. If a vulcanized soft-rubber ball at 20° C. be dropped from a known height onto the flat surface of a heavy steel plate, it will rebound about 77% of the distance dropped. If the same ball be cooled to −20° C., the rebound is only 24%. If cooled further to −40° C., the rebound is found to be 4%, but at −100° C., it has risen to about 70%. If heated to 100° C., the rebound is again high, about 83%. The changes in Young's modulus of elasticity of this rubber with temperature are no less instructive.* Over the temperature range 0° to 100° C., temperature has but little effect; the value averages 7000 lbs. per sq. in. However, dropping the temperature to −78° C. raises the modulus to about 1,200,000, an increase of 170-fold. It is obvious that these relatively limited changes in temperature have produced extraordinarily important changes in the elastic behavior of the rubber mass.†

Keeping in mind the physical arrangement of the molecules in unvulcanized rubber described on p. 407 *et seq.*, together with the fact that the change of major significance which occurs on vulcanization is the tying together of the chains at certain of their points of contact (p. 419), it is clear that any large distortion of the rubber mass involves slip of the molecules over each other, which in turn generates friction. This friction will be the greater, the greater the intermolecular attractions. The hysteresis loss which represents this internal friction will be increased by any factors which augment the van der Waals forces, *e.g.*, by drop in temperature with its accompanying con-

---

* The modulus is influenced by so many factors that it is difficult both to measure and to define. Thus, its value on release of load is less than on application of load, due to hysteresis. The quantity here discussed is the modulus under freshly applied, low loads:
† Both the temperature level of minimum resilience and the coefficients vary considerably with the stock.

tractions. The very small rebound at −40° C. means that only a minor fraction of the kinetic energy of the ball as it hits the plate is restored as such on rebound. The remainder must have dissipated itself as heat of internal friction within the ball; *i.e.*, at this low temperature level the molecules are close enough so that the partial valence attractive forces induce excessive friction as the molecular chains slip over each other.

At higher temperatures, the rubber expands, the intermolecular attractions decrease, energy lost by friction of slippage is reduced, and rebound increases. By 100° C., the hysteresis loss has become small.

At −100° C., on the other hand, the rubber molecules have been brought so closely together by the contraction caused by cooling that the mutual attractions between them at their points of contact are very great. This induces friction so high that any normal distorting force is inadequate to overcome static friction. The mass therefore behaves as though the molecules were cross-linked at all points of contact, so that they do not even start to slip. This, of course, eliminates energy dissipation due to friction, so that restitution must again be high.

At normal temperatures and above, while the friction due to intermolecular attractive forces and other resistances to distortion is sufficient to cause considerable dissipation of mechanical energy by hysteresis, it is not sufficient to have much effect on the distortion resulting from a given applied external stress. While there is a time lag in slip, the distortion equilibrium is ultimately reached. Since the restorative forces are not greatly influenced by temperature, this is equivalent to saying that Young's modulus is relatively independent of the temperature in this range.

However, at −100° C., where the attractive forces between molecules at substantially all points of contact have become sufficient so that under moderate stress they do not even start to slip, distortion must be due to deflection of valence bonds. This requires forces which are order of magnitude greater than those ordinarily necessary to cause slip of the rubber molecules over each other. In other words, Young's modulus, ratio of stress to strain, has increased inordinately; the ball is hard and

begins to approach steel in elastic behavior.* However, the development of this rigidity does not occur at a sharply defined temperature, but is extended over a temperature range. Thus, at −40° C., while there is considerable molecular slip accompanying rebound, it is doubtless true that at many of the points of contact between molecules the van der Waals attractive forces are sufficient even at this temperature level to prevent initiation of slip. Young's modulus starts to rise at temperatures well above those required for a high degree of rigidity, with its correspondingly high resilience.

It is a surprising fact that when one applies moderate tension loads (10 to 15% of the breaking load) up to a fixed maximum load to any given soft rubber stock, the hysteresis loop (p. 405) is remarkably insensitive to the rate at which the load is applied, provided the rate of loading and unloading is never excessive. Thus, a pure gum stock loaded to 300 lbs./sq. in. in 60 minutes gave a hysteresis loss of 25% of the work input, while for a duplicate sample similarly loaded in 1¾ minutes any increase in the loss was within the experimental error of perhaps 10%. The hysteresis is therefore not predominantly a viscous flow effect.† The energy degradation is evidently analogous to sliding friction, presumably of the molecular chains slipping over each other. One should note the difference in mechanism between this and viscous friction in fluid flow (p. 18). On the other hand, when in the test just described, one ceases to load and starts to reduce the tension (rates of loading and unloading being the same), elongation continues for a considerable time after load reduction starts. This fact cannot be explained on the single assumption of sliding friction independent of rate of slip, but indicates a component of viscous friction is playing a part in this phenomenon.

It is well recognized that the abrasive wear of a tire tread in actual road service is reduced at low temperatures. The above relationships make clear one reason for the difference. If a point on· the tread comes into contact with an obstacle in the

---

* Indeed, at this temperature level it must not be dropped too far or it will shatter.
† One might expect the rubber to have a definite equilibrium extension depending upon the load and to approach this extension by viscous flow in accordance with the formula of p. 24. The data indicate that this is not what happens, at least to any major degree.

form of a minor irregularity in the road surface, on the assumption that as the tread passes over the obstacle and starts to leave it the restitution of the rubber is complete, slippage between the surfaces of rubber and obstacle must occur. If, however, restitution is slow, not only is it possible to reduce the slippage, but the pressure of the tread on the surface of the obstacle and hence also the friction are lessened. Consequently, the high restitution characteristic of high temperature is sure to induce greater wear than the low restitution of a cold tread.

Vulcanization can also be effected by agents other than elemental sulfur. In the process known as cold-vulcanization rubber is vulcanized by dipping for a short time into a solution of sulfur monochloride in a suitable organic solvent. The reaction appears to be an addition of the chloride to the double bond:

$$\begin{array}{c}
\mathrm{CH_3{-}\overset{|}{C}} \\
\mathrm{\parallel} \\
\mathrm{H{-}\overset{|}{C}}
\end{array}
+ \mathrm{S_2Cl_2} +
\begin{array}{c}
\mathrm{\overset{|}{C}{-}CH_3} \\
\mathrm{\parallel} \\
\mathrm{\overset{|}{C}{-}H}
\end{array}
\rightarrow
\begin{array}{c}
\mathrm{CH_3{-}\overset{|}{C}{-}Cl \quad Cl{-}\overset{|}{C}{-}CH_3} \\
\mathrm{H{-}\overset{|}{C} \;-\; S \;-\; \overset{|}{C}{-}H}
\end{array}
+ \mathrm{S}$$

$$\mathrm{OR} \rightarrow
\begin{array}{c}
\mathrm{CH_3{-}\overset{|}{C}{-}Cl \quad Cl{-}\overset{|}{C}{-}CH_3} \\
\mathrm{H{-}\overset{|}{C} \;-\; S \;-\; \overset{|}{C}{-}H} \\
\mathrm{\parallel}
\end{array}$$
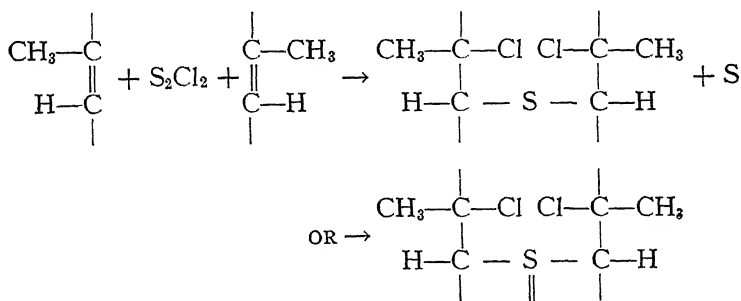
Since reaction is rapid, to prevent over-vulcanization of the external parts, the process is used only for thin articles. Products of physical properties superior to those of raw rubber can also be obtained by vulcanization with selenium and tellurium, although these elements do not give any products similar to hard rubber or ebonite. Presumably the mechanism of vulcanization in these cases is similar to that in the case of sulfur.

Vulcanization is almost always carried out at substantially constant temperature.* The rate of sulfur addition varies enormously (see below) but reaction kinetics are normal in type. Other things being equal, the rate is proportional to free sulfur content and hence falls off to low values as sulfur is consumed.

* When sulfur addition is very rapid, the heat of reaction may raise the temperature of the interior of the mass.

There is no indication of reversibility of the reaction under any normal conditions of cure. The optimum in physical properties with time of cure, despite continuing sulfur combination, demonstrates that the changes in the rubber are not limited to simple sulfur addition.

Figures 11 and 14 show that tensile strength passes through a maximum, followed by a more or less rapid decline. Study of the data shows that in all cases the initial rise in tensile strength is the more rapid the faster the rate of sulfur addition during that part of the cure; rapid drop off in tensile strength beyond the maximum is usually found in mixes in which further addition of sulfur during this period is small. These facts point convincingly to the conclusion that, whereas addition of sulfur tends to build up tensile strength (or other desirable physical properties), some other change is going on under the conditions of cure, which tends to neutralize and even undo the physical effects of sulfur addition. The second factor appears to be a thermal breakdown of vulcanized rubber at the temperatures employed for cure. Indeed, the disintegration of vulcanized rubber when exposed to moderate temperatures for relatively short times is one of its most serious weaknesses as a material of construction. On the basis of this assumption, one can explain the changes in properties of rubber during thermal cure as due to a great improvement induced by sulfur addition, accompanied by a steady breakdown due to thermal decomposition of the vulcanizate. During this period of breakdown unsaturation continues to diminish, indicating that the underlying change is chemical in character.

FIG. 11. Variation in Tensile Strength with Time of Cure.* 1, Rubber 100, Sulfur 1; 2, Rubber 100, Sulfur 1, Zinc oxide 1, Polybutyraldehyde aniline 0.5; 3, Rubber 100, Sulfur 3; 4, Rubber 100, Sulfur 3, Zinc oxide 5, Mercaptobenzothiazol 0.5; 5, Rubber 100, Sulfur 10. Temperature of cure, 142° C.

* Data of GARVEY, B. S., and WHITE, W. D., *Ind. Eng. Chem.* 25, 1042 (1933), and of GARVEY, B. S., and THOMPSON, G., *Ibid.* 25, 1292 (1933).

Although the problem of rubber vulcanization is extraordinarily complex, one thing is certain: the physical properties of a given mix * are not determined by the coefficient of vulcanization alone. A second important variable is time of heat during cure.† While study of the data indicates that these two are the major independent variables, other complicating factors are undoubtedly active and the man in the industry must be on the watch for their effects.

100 Smoked Sheets
11.11 Sulfur
Temperature 140°C.

o Unpreheated
× Preheated 20 min.
□ Preheated 120 min.
▾ Preheated 300 min.

Tensile Strength lb./i
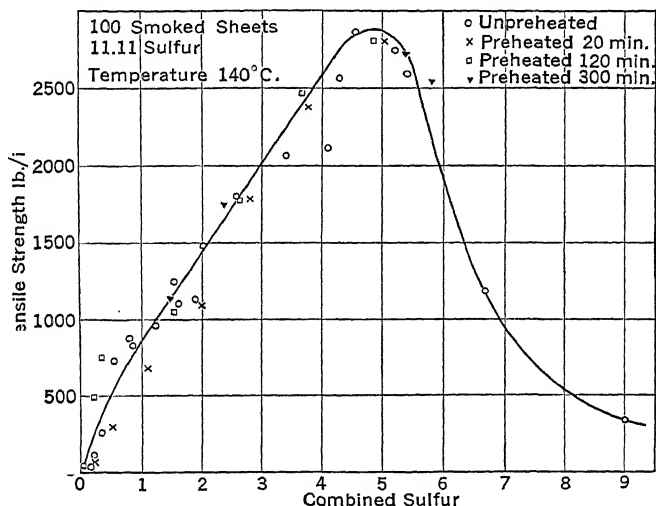
Combined Sulfur

FIG. 12.   Effect of Preheating upon Tensile Strength.‡

Accelerators.

Forty years ago soft rubber cures required upward of 8% of sulfur (always based on 100 parts of raw rubber), of which more than half remained uncombined.§   Even so, time of cure was a matter of hours, with consequent low plant capacity and deterioration during cure.   Clearly, catalysis of the reaction was needed.   Development of a series of suitable organic materials, the so-called accelerators, constitutes probably the most impor-

* Ignoring variations in minor constituents, e.g., in nature or quantity of accelerator or in initial sulfur.

† That heating prior to cure does not have the deleterious effect of heating during cure is shown by Fig. 12.   Evidently the vulcanizate is more sensitive to heat than the original rubber linkages.   It may be that sulfur bridges are thermally less stable than carbon bonds.

‡ NUTTING, R. D., Mass. Inst. Tech. Thesis (1933).

§ After cure some of the free sulfur diffused to the surface and crystallized as "bloom."

tant advance of the industry since the original discovery of vulcanization by Goodyear.

Comparison of curves 2 and 4 with 1, 3, and 5 of Figs. 11 and 13 indicates the magnitude of the result achieved. It will be noted that the time required for reaching high tensile strength is reduced even more rapidly than vulcanization rate is increased.* Accelerators also improve the physical properties of the product, not only because of reduction in curing time but also because of the possibility of lowering curing temperature. Thermal breakdown of vulcanized rubber appears to have a high temperature coefficient. Below about 140° C. it rapidly decreases, whereas at temperatures somewhat above this it becomes excessive. Keeping down the temperature of cure makes it possible to conserve the ultimate quality of the vulcanized rubber. The use of accelerators has also led to more uniform products and greater resistance to deterioration.

FIG. 13. Variation in Coefficient of Vulcanization with Time of Cure.† (For key to figure see Fig. 11.)

Accelerators, although frequently nitrogen- or sulfur-containing compounds, vary widely in their chemical nature. Consequently it is customary to classify them with regard to their speed rather than their constitution. Very fast accelerators, which act at temperatures below 100° C. in periods of a few minutes, are called ultra-accelerators; rapid accelerators have a vulcanization period of approximately thirty minutes at 100 to 140° C., while moderate and slowly acting accelerators take still longer times of cure to develop optimum mechanical properties at 140 to 150° C. Figure 14 shows the action of a number of characteristic accelerators upon the curing curve at 141° C. of a

* In industrial practice, the quality of the vulcanizate is always evaluated by physical tests, e.g., tensile strength at break, load required to give 500% elongation (the so-called modulus), etc., rather than by the coefficient of vulcanization, which is seldom determined, since (as pointed out above) it is not necessarily a guide to quality.

† Data of GARVEY, B. S., and WHITE, W. D., *Ind. Eng. Chem.* **25**, 1042 (1933), and of GARVEY, B. S., and THOMPSON, G., *Ibid.* **25**, 1292 (1933).

mix containing 100 parts rubber, 10 of zinc oxide, 2 of stearic acid, and varying amounts of sulfur. It will be seen that some accelerators tend to give sharp maxima, while others give a plateau effect, the tensile strength of the vulcanized mix changing but little with overcure. Very rapid accelerators have to be watched with care on the mill, since if the temperature rises too greatly, scorching or premature vulcanization occurs and the batch may be completely spoiled. The choice of an accelerator or mixture of accelerators depends largely upon the type of rubber goods being manufactured. Thus, for thick articles where rise of temperature in the interior will be slow, the use of an accelerator such as mercaptobenzothiazol, with which the optimum mechanical properties drop off but slowly with overcure (the so-called plateau effect), is indicated.

FIG. 14. Accelerators.* 1, Tetraethylthiuram disulfide 0.375, Sulfur 2.0; 2, Zinc diethyldithiocarbamate 0.375, Sulfur 2.0; 3, Butyraldehyde aniline 0.5, Sulfur 2.5; 4, Mercaptobenzothiazol 0.625, Sulfur 2.5; 5, Diphenylguanidine 1.0, Sulfur 3.0.

For modern, high-grade rubbers, active accelerators and low-sulfur mixes (0.5 to 1%) are used.

The variety in chemical nature of accelerators makes any common theory as to their action difficult. Possibly the simplest explanation is the postulation of an intermediate compound between sulfur and accelerator. This receives support from the fact that certain accelerators, such as tetraethylthiuram disulfide, can vulcanize rubber, even in the absence of free sulfur, giving up some of their own sulfur in the process. The changes may be represented:

$$S + A = AS$$
$$R + AS = RS + A.$$

* JONES, MALDYN, *Trans. Inst. Rubber Ind.* 11, 37 (1935)

It is likely that different accelerators have different mechanisms, the essential feature being liberation of sulfur in an active form for reaction with the rubber.

Any satisfactory theory must account for the much greater effectiveness of most accelerators in the presence of certain inorganic oxides (themselves without appreciable catalytic effect), notably zinc oxide. For example, with a stock containing 3 parts sulfur and 0.5 part mercaptobenzothiazol per 100 parts rubber, the tensile strength at break of the cured sample increased from about 92 kg. per cm.^{2} to 204 kg. per cm.^{2} upon incorporation in the mix of 5 parts zinc oxide. These materials are often called secondary accelerators.* The necessity for their use still remains unexplained.

Fillers and Reinforcing Agents.

As in the case of other plastics, an important function of rubber is as a bonding agent. The incorporation of suitable solids in a rubber mix can very greatly improve its physical properties, increasing not only its stiffness but also its tensile strength and resistance to abrasive wear. Such materials, of which carbon black is outstanding, are called reinforcing agents. In addition, other solids are frequently used, the so-called fillers. These serve merely as diluents for cheaper rubber goods, weakening the rubber (Table II). Coloring pigments may be added during the compounding process on the mill, although in some cases the pigment may also act to a minor degree as a reinforcing agent and vice versa.

The physical properties of the cured rubber are, in general, better the finer the particle size of the material added. The action probably parallels its effect in paints (p. 326). It has been suggested that carbon black reinforces by cross-bridging between rubber molecules (presumably by partial valence forces), much as does sulfur. However, it is difficult to see why its effect should be slight until sulfur vulcanization occurs or why

* Since the earliest days of the industry oxides have been used in rubber mixes, particularly litharge, magnesia, and lime. It seems probable that, in the absence of synthetic organic accelerators, these function as secondary accelerators of natural accelerators, perhaps proteins or resins, present in the raw rubber itself. Thus, the more highly natural rubber is purified, the slower its rate of vulcanization and the less satisfactory its cure. Synthetic rubbers cannot be vulcanized in absence of accelerators.

TABLE II.*   EFFECT OF LOADING (23 VOLUME PER CENT)
WITH VARIOUS POWDERS ON STRESS-STRAIN
PROPERTIES OF A RUBBER MIX †

Solid	Tensile Strength at Break, in lbs./sq. in.	Elongation at Break, in in.	Tensile Strength at Break (actual cross section), in lbs./sq. in.
Control	3350	7.75	26,000
Magnesium carbonate	4220	7.23	30,600
Zinc oxide	3970	7.17	28,500
Limestone whiting	2650	7.00	18,600
Talc	2480	7.10	17,600
Blanc fixe	3460	6.62	23,000

magnesium carbonate should behave so similarly to carbon black.

The data of Barron and Cotton, recalculated and presented in Figs. 15 to 19, illustrate typically the effect of fillers in rubber. These measurements were all made on a single base mix, containing 3% by weight of sulfur, 5% zinc oxide, and 1.5% diphenylguanidine, expressed, not on the mix, but per 100 parts by weight of rubber, into which five fillers were incorporated in varying amount.   Each sample was subjected to optimum time of cure to get maximum tensile strength.   Figures 15 and 16 give the tensile strength and percentage elongation at break.‡ In view of the elongation there was a corresponding reduction in cross section of the samples.   Neglecting the change in volume with elongation,§ the tensile strength has been recomputed on the cross section at break and plotted in Fig. 17.

The complete stress-strain diagrams were measured for each sample and a typical group of these, for carbon black, is presented in Fig. 18, with the experimental points indicated.   Inspection shows that, while at the start elongation is large relative to stress, a point is reached far below the breaking load beyond which the elongation is substantially linear against the stress.

* DAVIS, C. C., and BLAKE, J. T., *Chemistry and Technology of Rubber*, p. 388, Reinhold Publishing Corp. (1937).
† Mix: rubber 100 parts, S 0.5, Se 0.5, mercaptobenzothiazol 0.5, and tetramethylthiuram disulfide 0.5.
‡ Experimental points have been omitted to avoid confusion, but as the deviation of the individual points from the curves is surprisingly small, they are substantially dependable.
§ The volume decreases 1% for an elongation of about 520%.

This is equivalent to saying that at high loads the samples follow Hooke's Law, although deviating widely from it at low loads. Furthermore, at high loads the stress-strain lines are remarkably parallel, despite some tendency toward flattening as the breaking point is approached. If one extrapolates the straight-line portions back to zero load, one obtains an extrapolated elongation E_0, such that at high loads the stress-strain diagram conforms to the equation:

$$E = E_0 + \alpha S.$$

There is a large deviation from this equation at low loads; however, this deviation rapidly fades away to a negligible value as the load is increased. If one will employ the term,

$$E_0 e^{-0.00011\ E_0 S},$$

for this correction, the computed values of the elongation check the measurements of Barron and Cotton satisfactorily. The curves drawn on Fig. 18 are constructed from this equation. Similar curves fit the data for the other fillers fully as well as these for carbon black.
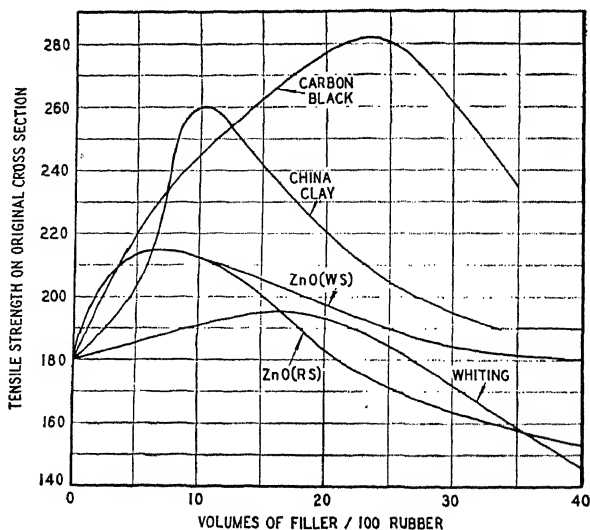
FIG. 15.

FIGS. 15–19.   Effects of Fillers on Elastic Properties of Rubber.*

CARBON BLACK

ZnO (WS)

CHINA CLAY

ZnO(R S)

WHITING

20

ES OF FILL

Fig. 17

TENSILE STRENGTH ON CROSS SECTION AT BREAK

ZnO (R.S)

ZnO (W.S)

WHITING

CHINA CLAY

CARBON BLACK

20   30

% LOADING

Fig. 19.

EXTRAPOLATED ELONGATION, E, %

100

2

1

ZnO(W)

ZnO (R)

WHITING

CARBON BLACK

CLAY

F FILL

Fig.

% ELONGATION AT BREAK

BASE MIX

8 VOLS. CARBON BLACK

23 VOLS.

34 VOLS.

150

SS,

% ELONGATION
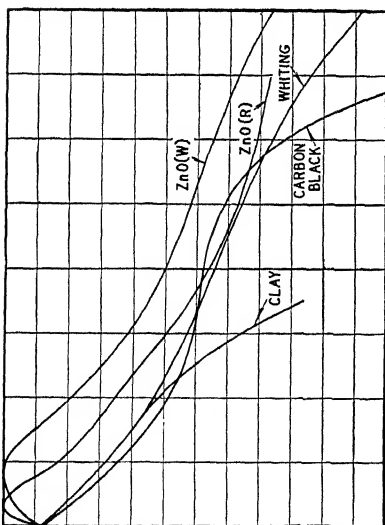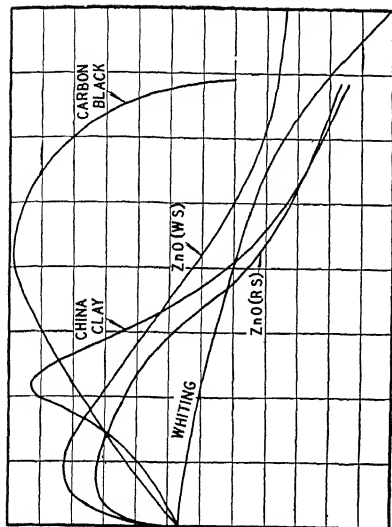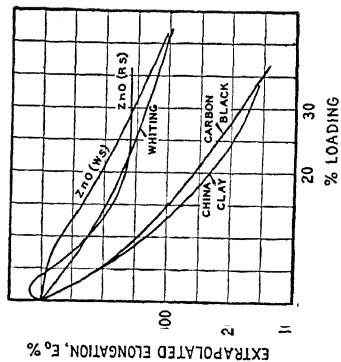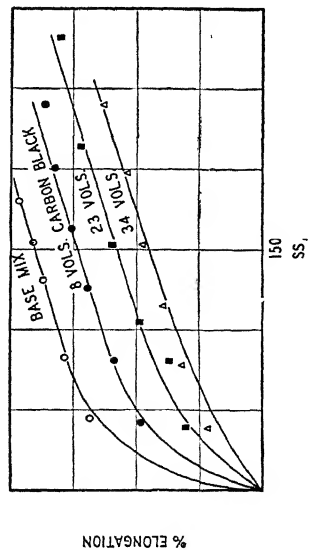
In other words, the distensibilities of samples prepared from this base mix depend only on the amount of filler in the mixture and a single quantity E_0, which is determined by the amount of filler but varies from one filler to another. The effect of the amounts of the different fillers in the mix on this extrapolated elongation E_0 is given by Fig. 19. Keeping in mind the fact that a high value of E_0 means a very low value of the modulus of elasticity at low loads, *i.e.*, a large elongation for a given low tensile stress, Fig. 19 shows that carbon black and china clay stiffen the rubber extraordinarily under low stress, even for relatively small per cent of the fillers. Zinc oxide at low percentages stiffens it not at all; indeed, one sample* shows a definite increase in E_0. At low percentages, whiting is intermediate between the zinc oxides and the other two loading materials, while at high percentage loading the zinc oxides and the whiting have approximately equal effect, all stiffening the samples far less than the carbon black and china clay.

Figures 15 and 17 show that all five loading materials increase tensile strength on original cross section to a maximum and all except whiting do the same for tensile strength on cross section at break. On the other hand, the maxima are reached at very low values of the filling agent in the case of zinc oxide, at somewhat higher values for china clay, but at the highest values of all for carbon black. Carbon black is the only pigment which gives a large stiffening effect at low stresses, as measured by low values of E_0, combined with the possibility of incorporating large quantities of it into the rubber before reaching the point of maximum tensile strength.

It is well to compare these relations with the data of Fig. 16. It is surprising what little effect the character of the loading agent has on elongation at break when present in moderate amount. The case of zinc oxide is an exception. Moreover, the zinc oxide curves of Fig. 16 parallel the others, except for the initial rise. This last effect may perhaps be due to the action of the zinc oxide on the vulcanization reactions themselves, perhaps modifying the mechanism of acceleration (see p. 430).

---

* The two zinc oxides used are known by the trade names red seal and white seal.

It is important to have the clearest possible interpretation of these relationships. To begin with, up to the point of maximum strength the volume per cent of filler is so low that the rubber is almost certainly a continuous external phase around the pigment particles, a conclusion fully confirmed by the similarity in elastic behavior of the filled stocks and the pure gum. This means that all the tension load must be carried by the rubber itself; hence, the increase in tensile strength resulting from incorporation of filler proves that the presence of the particles of filler makes the rubber itself around them stronger.

Since the great elongation of the stock occurring before rupture must involve during stretch a very large separation in the direction of stretch of adjacent pigment particles, the stiffening effect of the fillers indicates a stiffening of the rubber matrix itself rather than a mere interference with its movement around the particles. Any attempt to visualize the mechanism of filler action must be made in the light of these facts.

The flow of the rubber around the ultimate particles of filler during incorporation on the mill tends to orient the molecular chains of the rubber parallel to the surfaces of the particles of filler with which they are in immediate contact.* Where the volume per cent of filler is considerable and the particle size small, the distance from the surface of one particle of filler to that of its neighbor can be small relative to the average length of the rubber molecular chain so that a considerable percentage of the rubber molecules may be not only oriented parallel to the particle surfaces but have one end in contact with the surface of one particle and the other with that of its neighbor. If one will assume strong adhesion between the rubber molecule and the surface of the particle of filler, the partial orientation of the molecular chains must result in a considerable stiffening of the mass, because a significant percentage of the rubber molecules is already combed out in the direction of elongation before it starts. Furthermore, if the adhesion is sufficient, loading may well increase the breaking strength, because the unoriented molecular chains of unfilled rubber can rupture due to localized

* Molecular orientation is probably quite random parallel to particle surface, but near the surface there are few if any molecular chains whose axes are perpendicular to the surface.

overload resulting from snarling and kinking on elongation at gross loads less than those required to rupture the more oriented molecules of the filled rubber.*   The structure of a loaded stock

FIG. 20.   Diagrammatic Sketch of Structure of Loaded Rubber Stock.

may be indicated, highly diagrammatically, by Fig. 20, in which the first sketch shows the relative positions of a number of particles of filler before stretch, and the second, the corresponding positions of the same particles after stretch, the rubber chains effective in carrying the load being indicated by connecting lines.†   Granting this picture of structure, the parallelisms in behavior on elongations below the breaking load of the differ-
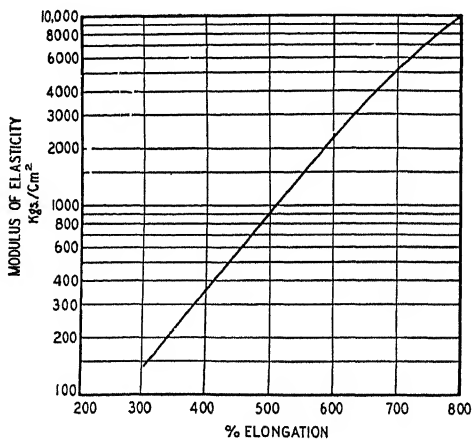
FIG. 21.   Effect of Distortion on Young's Modulus for Loaded Rubber.

ent fillers of Fig. 19 are due to the fact that the differences in particle size are not marked; the differences in breaking load (Figs. 15 and 17) would be ascribed to adhesion, *i.e.*, adhesion

* The unloaded rubber might rupture by a slippage of the molecules past each other which could not occur with the loaded rubber if adhesion between filler and rubber is sufficient.   However, such slippage ought not to be the major cause of rupture if three-dimensional cross-linkage by vulcanization is adequate.

† Before stretch starts most of the chains must be far from straight, a point difficult to indicate in the diagram.

of rubber to carbon black is far greater than its adhesion to china clay; the stiffening which accompanies elongation (Fig. 21) is caused by the progressive pulling out of the carbon chains.

Manufacture from Latex.

It is possible to prepare many classes of rubber goods, not readily made from sheet rubber, directly from preserved rubber latex (p. 402), thus omitting the coagulation and mastication steps necessary in the usual procedure. Due to lack of breakdown, latex can also give increased strength. Because it is many times less viscous than a solution of masticated rubber in benzene (or similar solvent) of equal concentration, latex is used chiefly to replace such rubber solutions or cements. In the case of latex, the "solvent" (water) is non-poisonous, non-inflammable, and does not require recovery (cf., lacquer emulsions, p. 322). Furthermore, no expensive machinery is needed for compounding.

In general, latex, first compounded with fillers, accelerators, sulfur, etc., is given the shape of the finished article by some technique of gelation and is finally vulcanized to render the shape permanent. Because many powders are positively charged, tending to coagulate the negatively charged latex,* it is usually necessary before addition to the latex to disperse them in water [often with the aid of a colloid mill (p. 113) and a dispersing agent, e.g., sodium naphthenate], and to add a protective colloid, such as glue or casein. During dispersion too vigorous mechanical agitation must be avoided, since this ultimately causes coagulation.

By compounding with zinc oxide and an ultra-accelerator, followed by heating at 75 to 100° C. for a sufficient time, latex may be vulcanized without loss of water so that the individual particles remain dispersed and no coagulation occurs. On drying, however, the individual particles in the residual film, despite their vulcanization, form a continuous sheet of rubber. The mechanism of coalescence is uncertain. It has been suggested that it is due to interaction of unsaturated sulfur bonds on the surface of the vulcanized latex particles.

* Even if uncharged, when added dry their abstraction of water is frequently sufficient to cause local coagulation.

Formation of the rubber article from latex may be accomplished by a variety of methods. Balloons, rubber gloves, certain kinds of footwear, etc., are often made by dipping forms of the desired shape into compounded latex, drying and then vulcanizing by heating in hot air or water. Thicker films can be obtained by dipping several times or, more ingeniously, by adding thixotropic materials such as bentonite (p. 243), so that the mixture gels, leaving no tendency for the film to thin by drainage on pulling out the form. Casting may also be used to form the article. Certain salts, ammonium nitrate, sodium silicofluoride, etc., can be added to latex without causing gelation in the cold, but producing a sensitized dispersion, which on warming may be gelled in molds. To secure uniformity or, in hollow articles, evenness of wall, the molds are rotated during the simultaneous gelation and vulcanization process. Electro-deposition, using an anode of the desired shape, can also be used. The negatively charged rubber particles travel towards the anode, as do the fillers and vulcanizing ingredients, forming a uniform film which can be detached and vulcanized. The fact that both compounded solids and rubber particles travel towards the anode at the same speed may be accounted for by assuming that all are covered with an adsorbed film of the same protective colloid (cf., collodion particles and gelatin, p. 178). Current continues to flow even after comparatively thick films have been built up, sufficient liquid being left in the interstices to carry the current. Indeed, the particles are uncoagulated; on reversing the current they move back into solution.

Rubber thread can be obtained from latex by spinning it into a coagulating bath, containing ammonium acetate and acetic acid, from glass jets (cf. viscose rayon, p. 364), subsequently passing the thread through hot glycerol baths to dry and vulcanize it. Sensitized latex (*vide infra*) may also be used, being spun into warm air, where it at once gels (cf. cellulose acetate rayon, p. 375) and can be dried and vulcanized. Latex can also be used for impregnating paper, textiles, leather, etc., producing materials useful for many purposes. In these products it may function both as bonding agent and as an impregnant, *e.g.*, for waterproofing.

Ageing.

Vulcanized rubber is, unfortunately, subject to deterioration in many of its physical properties during storage and use.   In general, such changes, known as "ageing," are characterized by a gradual hardening and increase in brittleness of the rubber, accompanied by loss of tensile strength and elongation.   The changes are catalyzed by heat, light, and the presence of traces of such metals as copper and manganese.*   Undoubtedly ageing is due to oxygen attack of the rubber, since oxygen content goes up and the rubber can be kept almost indefinitely if oxygen be excluded.

Because it is important to predict how rubber will age over considerable time periods, many attempts have been made to develop rapid or dependable accelerated ageing tests.   Two only need be mentioned: the Geer oven test and the oxygen bomb test, the latter apparently being the more reliable.   In the former, the test pieces are hung in an oven at 70° C., with air circulating around the specimens; in this way, attack by oxygen is accelerated.   Similarly, in the bomb test, attack is accelerated by using oxygen at 300 lbs. per sq. in. pressure and warming the bomb to 70° C., although lower temperatures are sometimes preferred.   Deterioration is usually judged by appearance, decrease in tensile strength, and the like.

The factors governing ageing are very complex and frequently baffling.   Increase in temperature rapidly accelerates deterioration but high relative humidity retards it.   Light, particularly ultraviolet, also speeds up ageing, although normally this is predominantly a surface effect.   Excessive combined sulfur accelerates ageing, although the deterioration correlates more closely with overcure (in the sense that physical properties have passed the optimum), than with any specific value of combined sulfur itself.   Consequently, it is considered good practice to stop cure short of the optimum point.   Before the widespread use of accelerators, stocks frequently became tacky in the earlier stages of ageing, getting brittle later, suggesting the possibility of two distinct mechanisms of oxygen attack, one resulting in

* Deterioration can occur due to heat alone, e.g., in high-pressure steam hose.

broken linkages, as in milling, and the other giving intermolecular bridging as in sulfur vulcanization. With modern low-sulfur cures tackiness seldom develops.

An important development of recent years has been the discovery of a series of anti-oxidants which inhibit the oxidation of vulcanized rubber, greatly improving its ageing properties. These are generally hydroxy compounds or bases, *e.g.*, *p*-hydroxydiphenyl, phenyl β-naphthylamine, etc. The rate of absorption of oxygen by rubber is autocatalytic, initially slow but becoming progressively more rapid as oxygen is absorbed. These anti-oxidants generally prolong the inhibition period, *i.e.*, retard the onset of oxygen absorption by the stock without greatly affecting the nature or rate of the action when once it sets in. It may be that the initial attack is a slow formation of peroxide linkages, the excess oxygen of which is taken up preferentially by the anti-oxidants, preventing its further destructive combination with the rubber until the anti-oxidants themselves have been consumed.

Reclaimed Rubber.

As already stated, if low-sulfur mixes are overcured they deteriorate, with loss in tensile strength and other desirable physical properties, the deterioration increasing rapidly with temperature. If heated sufficiently,* the rubber can again be reduced to a plastic mass. Practically all free sulfur will combine during this treatment, but since the initial sulfur is low, the product will still have a high degree of unsaturation, so that, if additional sulfur be incorporated, it will add under heat, again reducing plasticity and curing the stock. Advantage is taken of these facts in the so-called reclaiming of rubber, recovering it in a plastic state from factory scrap and discarded rubber articles. Reclaiming is complicated by (*a*) accompanying materials such as the wire beads of tires, and particularly the cellulose fabric of tire casings, hose, etc., (*b*) the compounding ingredients in the stock, and (*c*) the fact that the replasticized vulcanizate differs significantly from raw rubber.

* The result depends on both temperature and time under heat. Unfortunately, no method has so far been developed to protect rubber products from high temperature deterioration.

The stock is carefully sorted, shredded, and treated mechanically to remove foreign materials, as iron by magnetic action and shreds of fabric by cyclone-type separator. It is then heated to secure plasticization, the most important technique involving digestion in a 4 to 8% caustic soda solution for from eight to twenty-four hours, at temperature levels in the neighborhood of 180 to 200° C. (alkali reclaiming process). One major function of the caustic is to hydrolyze the cellulose, degenerating it sufficiently so that it can be removed satisfactorily by subsequent washing.

The transformation of the rubber is presumably a depolymerization of the vulcanizate.* As in milling, oxygen greatly accelerates the breakdown; thus, recovery can sometimes be effected by heating in live steam at atmospheric pressure, probably due to access of air under these conditions. Oxidizing agents, such as sodium hypochlorite, are sometimes added in the pressure process. Because the action is primarily depolymerization, it is aided by solvents, various oils being frequently introduced with the stock for this purpose. The term "devulcanization" is sometimes mistakenly applied to the process; but, far from removing any of the combined sulfur of the original stock, combined sulfur increases at the expense of the free sulfur, despite the tendency of caustic to react with the latter.

The plasticized rubber is washed free of excess caustic, dewatered by squeezing (e.g., in a cone press) to 50% of water or somewhat less, and dried. High gas temperatures can be used in the early stages of drying, because the stock is protected against undue temperature rise by the wet-bulb effect. The stock is milled on so-called refiners, rolls closely set so that nonplastic particles of whatever sort will be held in the bank. The stock is usually extruded through strainers, in which it is forced through screens under pressure to eliminate lumpy foreign matter. It is finally sheeted for shipment or ultimate use.

Reclaim is more plastic than ordinary rubber and has far less resilience or nerve, so that it acts as a softener in raw rubber mixes and is extremely valuable for use in extrusion operations,

---

* Unfortunately, nothing certain is known of the mechanism of molecular disintegration or of the structure of the product.

It likewise cures more rapidly than ordinary rubber. However, it vulcanizes to a product inferior in physical properties, the deterioration increasing with successive reclamations.* It is almost always used admixed with more or less new rubber. Because the compounding ingredients are not removed by reclaiming, allowance must be made for them in using the product.

Artificial Rubbers.

It was early discovered that isoprene, under the influence of heat, hydrochloric acid, or sodium, or even if allowed to stand for a long period of time, polymerized to give an elastic body which possessed many of the physical characteristics of rubber. Most of the substances possessing physical properties similar to those of natural rubber, which have found use as "synthetic" or artificial rubber, are polymers of butadiene or its derivatives. In general, polymerizability appears to decrease with increased substitution and is favored by the presence of unsubstituted, terminal hydrogen atoms.

Recent years have seen the introduction of an important group of thermoplastic products of high molecular weight, presumably linear polymers, all of which are distinctly rubber-like in physical characteristics. Of these, the polymer of isobutylene may be chosen as typical, because of its ease of preparation and simplicity of chemical structure. If a solution of isobutylene in a volatile solvent such as ethylene be treated with BF_3 as a catalyst, it polymerizes practically instantaneously. The reaction is evidently of the chain type, each successive molecule adding to the chain activating its neighbor. Low temperature of formation favors high molecular weight. Continued treatment with catalyst produces no further polymerization. The product has a single residual double bond per molecule, and is in consequence highly resistant to oxidation.† The exact structure is uncertain, but the disappearance of a double bond for each

* On this account the sorting operation before reclaiming is unusually important. While articles originally compounded with considerable reclaim can again be subjected to the process, the product from it will depress the quality of vulcanizates into which it is incorporated. However, such material can be very useful in articles in which manufacturing cost must be low but in which strength, resilience, etc., of the rubber itself is of minor importance, such as insulated wires.

† Thus its incorporation in rubber mixes can markedly reduce deterioration or ageing due to oxidation.

reacting monomer but one indicates a type such as

$$\underset{\underset{H}{\overset{CH_3}{|}}}{-CH_2-C} - CH_2-CH_2-\underset{\underset{H}{\overset{CH_3}{|}}}{C} - CH_2- \quad \text{OR} \quad -\underset{\underset{CH_3}{\overset{CH_2}{|}}}{C}-CH_2-\underset{\underset{CH_3}{\overset{CH_3}{|}}}{C}-CH_2$$
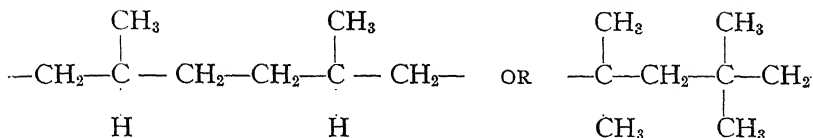
The polymer dissociates under sufficient heat largely to the original isobutylene. It is available commercially under the trade name vistanex in grades differing predominantly in chain length, and is used widely, *e.g.*, as an addition agent to oils to increase viscosity and particularly to decrease its temperature coefficient.
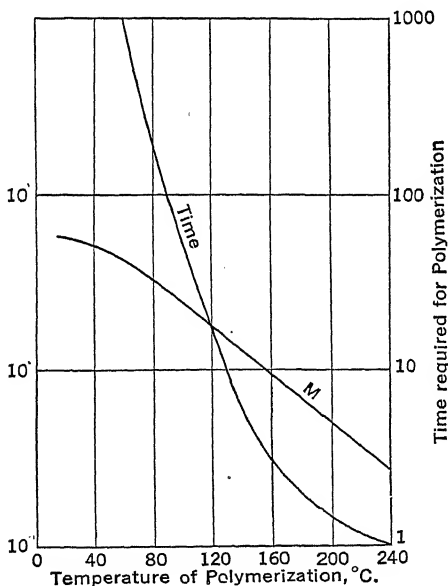
FIG. 22. Styrene Polymerization.*

The nature of the reactions involved in cases of this type is clarified by study of Fig. 22, which shows the time required for thermal polymerization of liquid styrene and the mean molecular weight of the polymer produced, both as functions of the temperature of polymerization. Inspection shows that reaction rate

* Based on data of SCHULZ, G. V., and HUSEMANN. E., *Z. physik. Chem.* B34, 187 (1936). On this plot time is in hours.

rises rapidly with the temperature, requiring approximately 12° C. to double at low temperatures but far more at high.   The molecular weight of the product decreases rapidly as temperature rises, halving for a rise of approximately 45° C., except at low temperatures.   The molecular weight is apparently not greatly affected by the extent of polymerization (*i.e.*, percentage of conversion) of the monomer.   The curves suggest the possibility that at low temperatures nuclei of polymerization (*i.e.*, molecules activated sufficiently to initiate reaction) are formed at a relatively low rate.   On the other hand, once reaction is made possible by the activation of a given molecule of monomer, its reaction with its immediate neighbor is sufficiently violent to activate in turn the double bond of the latter, so that polymerization goes on as a continuing chain reaction.   This phase evidently occurs at a very high rate.   Ultimately, on the other hand, the end group on the chain, however active, may not find a molecule of styrene sufficiently close to it to react with it before its activation energy has been dissipated.   This, of course, results in stoppage of chain growth.   The data indicate that stoppage occurs much sooner at high temperature than at low, conceivably due to the fact that at high temperature levels the styrene molecules are farther from each other on the average as shown by the lower density of the liquid.

Ethylene dichloride reacts with an aqueous solution of sodium polysulfide progressively, with elimination of the chlorine from the organic molecule:

$$ClH_2C.CH_2Cl + Na_2S_4 \rightarrow ClH_2C.CH_2S_4Na + NaCl,$$

$$ClH_2C.CH_2S_4Na + ClH_2C.CH_2S_4Na$$
$$\rightarrow ClH_2C.CH_2.S_4.H_2C.CH_2S_4Na + NaCl.$$
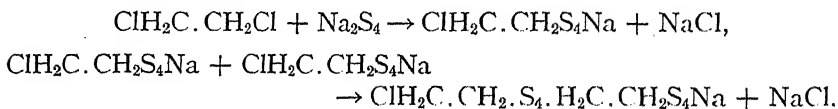
The result is a latex of the polymer, which, if formed under properly controlled conditions, is sufficiently dispersed so that effective washing for removal of the salts is possible.   The latex is coagulated with acid, milled, softeners and fillers are introduced, and it is molded and cured by heat.   This "vulcanization" involves no added sulfur but does require zinc oxide; it results in definite stiffening, but is thought to involve only some additional polymerization.   The product, the trade name of

which is thiokol, is especially valuable because of its resistance to swelling in solvents.

Another material is koroseal, the commercial polymer of vinyl chloride, more sensitive to solvents than thiokol, but particularly resistant to chemical action. A great deal of work on the syntheses of products of this general type from simple monomers is under way, and rapid progress along practical lines is to be anticipated.

A similar product is chlorinated rubber. Chlorine reacts with rubber readily, e.g., in carbon tetrachloride solution, but the reaction is by no means addition only, as large quantities of HCl are evolved. Neither is it limited to addition and substitution, since, depending on the conditions during certain steps of the reaction, disappearance of double bonds far exceeds total mols of chlorine combined. One must assume development of carbon to carbon linkages, as by elimination of chlorine from one carbon atom and hydrogen from another. This might result in cross-linkage between chains, but the solubility and viscosity characteristics of the products discredit such an assumption.* There is apparently cyclization along the chain but the rings break again as reaction proceeds. Intermediate products are unstable, splitting out HCl, but if the reaction is carried to 60 to 65% chlorine content, stability is satisfactory. The product is soluble in aromatic and chlorinated hydrocarbons, thermoplastic, non-inflammable, and remarkably resistant to aqueous acids and oxidizing agents. Thus, important uses are in corrosion-resisting parts, electrical insulations, etc.

Butadiene itself can be polymerized slowly, for example, by metallic sodium as catalyst, yielding so-called *buna* † rubber, one of the earliest attempted synthetic substitutes for the natural gum,‡ first produced in Germany. While the structure of the polymer is uncertain, there is some evidence of the following

* Viscosities of the solutions indicate reduction in chain length during the process.
† The name is made up of the first two letters of the monomer and the symbol of the catalyst.
‡ The fact that rubber itself is a derivative of isoprene has led to efforts to polymerize various conjugated di-olefines. This seems all the more attractive, because of the possibility of synthesizing these hydrocarbons from cheap raw materials, such as oil and coal. Unfortunately, yields are low and the problems of purification are formidable. While such sources will probably in the long run prove to be most satisfactory, other syntheses of the di-olefines required have been used heretofore.

reaction mechanism:

$$n\ CH_2\!\!=\!\!CH\!\!-\!\!CH\!\!=\!\!CH_2 \rightarrow$$
(1,3-butadiene)

$$
\begin{array}{ccc}
\underset{|}{H} & & \underset{|}{H} \\
-\!C\!-\!CH_2\!-\!C\!-\!CH_2\!- & \\
\underset{||}{CH} & & \underset{||}{CH} \\
CH_2 & & CH_2
\end{array}
$$
(buna rubber)
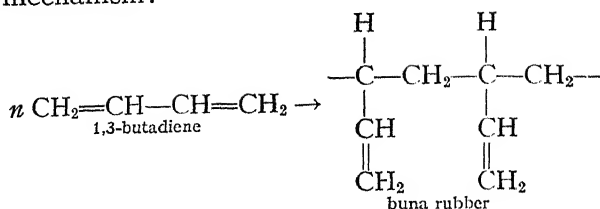
Butadiene has been synthesized from acetylene by (1) hydrolysis to acetaldehyde with a mercuric salt catalyst in the presence of dilute sulfuric acid, (2) condensation to aldol with dilute NaOH catalyst, (3) catalytic hydrogenation under pressure to 1,3-butylene glycol, and finally (4) dehydration to the hydrocarbon. In Russia butadiene is said to be prepared from ethyl alcohol by passing it over a special catalyst, resulting in the dehydrogenation of one molecule to acetaldehyde, followed by dehydration of this with a second alcohol:

$$CH_3CH_2OH + CH_3CHO \rightarrow CH_2\!\!=\!\!CH\!\!-\!\!CH\!\!=\!\!CH_2 + 2\ H_2O.$$
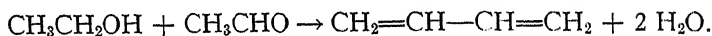
Polymerization may be effected not only with metallic sodium, but also with an emulsion of the butadiene in water, stabilized with the aid of a protective colloid and using peroxides as catalysts. In this way a latex resembling that of natural rubber can be obtained. The artificial rubbers thus prepared appear, from the limited experimental data available, to have somewhat low tensile strengths of the order of 30 to 50 kg. per sq. cm. upon vulcanization. These are, however, markedly improved by the incorporation of carbon black, which gives vulcanizates with tensile strengths of 180 to 220 kg. per sq. cm.

In 1918, 2,3-dimethylbutadiene, obtained by condensation of acetone, was polymerized in Germany as a rubber substitute, giving the so-called methyl rubber:

$$
n\ CH_2\!\!=\!\!\underset{\underset{CH_3}{|}}{C} - \underset{\underset{CH_3}{|}}{C}\!\!=\!\!CH_2 \rightarrow
$$

$$
-CH_2\!-\!\underset{\underset{CH_3}{|}}{C}=\underset{\underset{CH_3}{|}}{C}\!-\!CH_2\!-\!CH_2\!-\!\underset{\underset{CH_3}{|}}{C}=\underset{\underset{CH_3}{|}}{C}\!-\!CH_2\!-.
$$
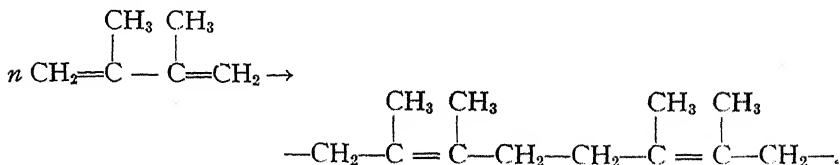
More recently a copolymer of butadiene and styrene (C_6H_5.-

CH : CH$_2$), known as buna-*S*, has been introduced, which exhibits excellent abrasive and heat resistance and is particularly suitable for tire treads. Butadiene also copolymerizes with acrylonitrile, CH$_2$=CH—C≡N, in aqueous emulsion to form a synthetic latex giving buna-N or perbunan; there are products of varying quality, presumably differing in ratio of the monomers and perhaps in the mechanism of polymerization. They are resistant to abrasion, heat, and the action of solvents, and age well, but are hard to mill and have low tensile strength. However, on incorporation of carbon black their strength increases to values exceeding those of rubber.

In the United States an artificial rubber has been developed by the polymerization of 2-chloro-1,3-butadiene, which is prepared from acetylene in the following stages:

$$2\,CH{\equiv}CH \xrightarrow[\text{NH}_4\text{Cl}]{\text{CuCl}} CH_2{=}CH.C{\equiv}CH \xrightarrow{\text{HCl}} Cl.CH_2.CH{=}C{=}CH_2$$

vinyl acetylene · · · · 4-chlor-1,2-butadiene

$$CH_2{=}\underset{\underset{Cl}{|}}{C}{-}\underset{\underset{H}{|}}{C}{=}CH_2$$

chloroprene monomer
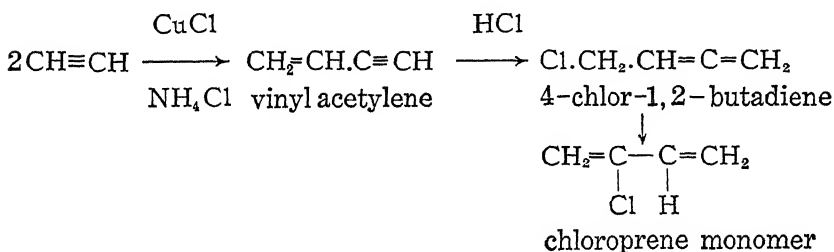
Chloroprene polymerizes much more readily than does isoprene, just as vinyl chloride (p. 480) polymerizes more readily than ethylene. On standing for ten days polymerization is complete, yielding a transparent, elastic substance which resembles soft vulcanized rubber in being non-plastic and unworkable on the roll mill. If, however, polymerization is carried only part way and the polymer is isolated from the unchanged chloroprene, an elastic body is obtained which can be worked on the rolls like natural rubber and on heating goes over into the non-plastic insoluble polymer. This polymer may be kept for several years in plastic condition at room temperature by the addition of a small amount of phenyl β-naphthylamine. Vulcanization * of neoprene (the trade name by which the polymer is now known)

---

* The word *vulcanization* is often used because the changes in physical properties under heat parallel those occurring in true vulcanization of rubber. The term *cure* is more appropriate (see p. 417).

therefore requires no sulfur, polymerization being simply com-
pleted by heating. The mechanical properties are enhanced,
however, by incorporation of metallic oxides, although carbon
black, in contrast to its effect on natural rubber and the other
synthetic rubbers, does not greatly increase its tensile strength.

Chloroprene can also be polymerized in emulsion form, so-
dium oleate acting both as a stabilizer and as a polymerization
catalyst, the reaction being complete in from two to eight hours
at room temperature. The latex corresponds to a vulcanized
natural latex, polymerization having proceeded to the final stage.

Neoprene compares favorably in strength with natural
rubber, but is far superior in its resistance towards mineral oils,
a fact which dictates many of its industrial uses. It ages well,
but is hard to mill, has poor heat resistance, and changes in
stiffness rapidly with temperature. Some of its physical proper-
ties, e.g., permanent set, can be improved by using a small
amount of sulfur during the vulcanization, thereby accelerating
the rate of vulcanization (polymerization) of the neoprene
as such.

BIBLIOGRAPHY

AM. SOC. FOR TESTING MATERIALS, Symposium on Rubber, March 9, 1932.
CRAMER, H. I., Ind. Eng. Chem., 34, 243 (1942).
DAVIS, C. C., and BLAKE, J. T., eds., Chemistry and Technology of Rubber,
    Reinhold (1937). (Contains a helpful bibliography.)
DAWSON, T. R., and PORRITT, B. D., Rubber, Physical and Chemical Prop-
    erties, Research Assoc. of Br. Rubber Manufacturers, England (1935).
    (Also contains a helpful bibliography.)
FISHER, H. L., Proc. A.S.T.M., 41, 521 (1941).
NOBLE, R. J., Latex in Industry, The Rubber Age (1936).
PORRITT, B. D., Chemistry of Rubber, Gurney and Jackson, London (1913).
SCHIDROWITZ, P., Rubber, D. Van Nostrand (1920).
WEBER, LOTHAR, Chemistry of Rubber Manufacture, Griffin, London (1926).
High Polymers: A series of monographs on the Chemistry, Physics, and
    Technology of High Polymeric Substances is being published by the
    Interscience Publishing Co., Inc., New York.

Chapter XIX

Ceramic Industries

The ceramic industries, including the manufacture of brick, stoneware, pottery, refractories, and the like, are based upon the following properties of clays:

1. The clays are plastic when wet with water, moldable into almost any form by various methods, and yet have sufficient strength (yield point) to retain the form imparted to them.
2. Evaporation of the water from the plastic clay (drying) completely eliminates its plasticity, materially increasing its mechanical strength, although the product again becomes plastic on wetting.
3. On heating to temperatures from 200° to 800° the clay decomposes progressively, evolving water of chemical constitution, leaving a relatively strong, porous mass which, however, is no longer capable of replasticization by water.
4. By further heating, a portion of the material may be melted, which will serve as a bond for the residual, unmolten material, leaving the shape substantially unchanged, provided the skeletal material is sufficient in amount.

Clay.

Clays are essentially hydrated silicates of aluminum, formed by natural hydrolysis or weathering of feldspar and various alkaline alumino-silicates making up granite and other igneous rocks. Because of the insolubility and inertness of the clays, their investigation and characterization is rendered difficult, and it was formerly supposed that there were a large number of

individual constituents, but recent work renders this doubtful. Altogether there appear to be about seven chemically distinct clay materials, of which kaolinite, $Al_2O_3.2\ SiO_2.2\ H_2O$, beidellite, $Al_2O_3.3\ SiO_2.H_2O$, and montmorillonite, $Al_2O_3.4\ SiO_2.H_2O$, are probably the most important. The first is the essential constituent of ball and china clays and many fire clays. Beidellite occurs widely in the finer fractions of clays made up predominantly of other materials. Montmorillonite is the principal constituent of bentonite and fuller's earth.* Because the natural clays usually consist of mixtures of the clay minerals, the ratio of alumina to silica can vary within wide limits. However, natural clays are in the majority of cases mixed with other minerals, such as undecomposed feldspar, silica, calcium carbonate, and the like, which do not possess clay-like properties.

The weathering by which the clays are synthesized in nature is a hydrolysis by atmospheric agencies, eliminating the alkalies by some such type reaction as the following:

$$K_2O.Al_2O_3.6\ SiO_2 + 2\ H_2O + CO_2 \rightarrow$$
$$Al_2O_3.2\ SiO_2.2\ H_2O + 4\ SiO_2 + K_2CO_3.$$
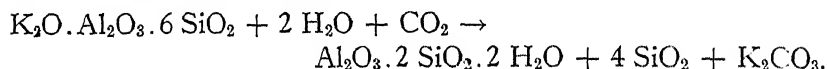
Clays remaining at the point of hydrolysis are called primary. Frequently, they have been transported by atmospheric agencies, particularly surface water, and deposited elsewhere in beds more or less stratified, as secondary clays. All shales belong in this class, likewise ball clays and frequently the fire clays. Secondary clays are generally more plastic than the primary and more apt to be contaminated with organic impurities, hydrated iron oxide, and the like.

In Table I are given the chemical analyses of typical potters' clays, together with that theoretically required by the formula $Al_2O_3.2\ SiO_2.2\ H_2O$ for kaolinite. It must be borne in mind that ball clays and fire clays are particularly variable in composition; thus, satisfactory fire clays range in silica content from about 40 to 65%. Ultimate chemical analysis is, therefore, an inadequate guide to the utility of a clay for the potter, except

---

* Today the term "fuller's earth" is restricted largely if not completely to clays of high plasticity which, when dried and ignited to high temperature levels insufficient to drive out all the chemically combined water, are effective adsorptive decolorizing agents for hydrocarbon as well as animal and vegetable oils. The same clays are almost always capable of functioning in the so-called contacting process for decolorization of mineral oils.

that it discloses undesirable constituents, such as iron, alkalies and alkaline earths. Iron discolors the finished ware and both alkalies and alkaline earths seriously lower the fusion point of the clay. The more plastic ball and fire clays generally contain 1 to 3% carbon, usually in the form of particles of lignite, as a consequence of which they are often blue or black in color.

TABLE I. ANALYSES OF TYPICAL CLAYS

	Kaolinite	China and Ball Clays	Fire Clay	Stoneware Clay
SiO_2	46.5	46	60	65
Al_2O_3	39.5	36	28	23
TiO_2		1		1
$Fe_2O_3 + FeO$		1		2
$CaO + MgO$		1		1
$K_2O + Na_2O$		1	1	1
Combined H_2O	14.0	14	10	7

The characteristics of clays which are of interest to the potter are determined, rather than by chemical composition, by the properties of the individual clay particles, most important of which are size, shape, and nature of surface. There is little doubt that plastic, clay-like properties are not developed in particles with sizes above about $10\ \mu$. The lower limit is more uncertain, although the existence of particles as small as $10\ m\mu$ has been reported. Particles larger than $10\ \mu$ are generally considered as impurities, silt, sand, etc. The shape of the smaller particles, below the resolving power of the microscope, is difficult to determine directly, although the larger particles appear as flat, plate-like crystals. That even the smallest particles are plate-like is indicated by the double refraction of their suspensions in viscous motion (p. 134). Pure clay minerals also exhibit marked X-ray patterns, characteristic of crystalline particles. Hence, it may be inferred that clays are composed largely of exceedingly fine, plate-like crystals. Recent X-ray work has shown that the plates themselves probably consist of alternate layers, as it were, of alumina and silica, bonded together through oxygen bridges. The chemically bound water is almost certainly present as hydroxyl groups, and is not liberated until fairly high temperatures are reached (Fig. 1). Simul-

taneously with release of the combined water the lattice structure of the clay mineral breaks down, leaving a mixture of alumina and silica.*

The hydrogens of at least some of the hydroxyl groups are mildly acidic in character, replaceable by basic ions. The natural clays usually contain alkali and alkaline earth metals
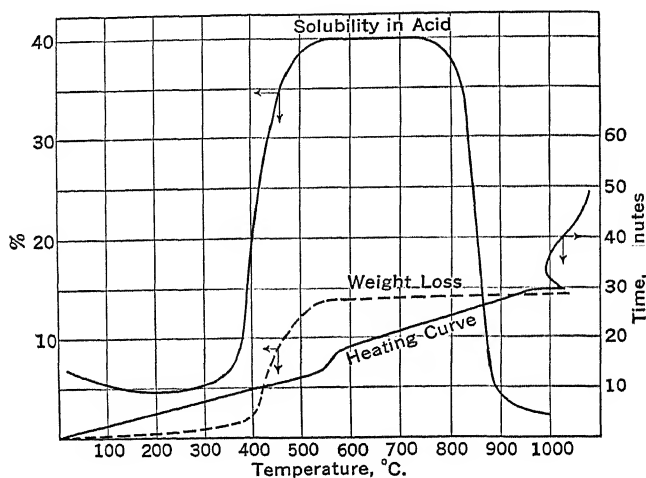
FIG. 1.   Ignition of Kaolin.

thus combined with them. These metallic ions may be re-moved by electrodialysis (p. 136), by washing with acid, or by double decomposition with the soluble salts of other metals. The first two methods produce the so-called acid or hydrogen clay; the last is known as base exchange, a phenomenon en-countered in many minerals besides the clays, such, for example, as the zeolites. Hydrogen clay on treatment with alkali behaves as a weak acid of high molecular weight. Its base-exchange capacity can be evaluated by measuring the milli-equivalents of base which can combine with 100 gms. of the acid clay, e.g., by titration equivalent to that of Fig. 8, p. 205; the two frac-tions of bentonite there tested had average equivalent spherical diameters (as determined by centrifugal sedimentation) of 14.3 mμ and 87 mμ respectively. The identity of the equivalent

* The recovery of alumina by acid digestion of such a thermally decomposed clay as indicated in Fig. 1, is an important industrial source of aluminum salts.

weights indicates either that the active hydrogens are independent of the amount of surface or that particle shapes are such that total surface is proportional to mass. In contrast to kaolinite, with a base-exchange capacity of about 10, that of bentonite is roughly 92. Kaolin is, however, a far weaker acid (thus, metal-free kaolin suspended in water has a pH of approximately 4 and bentonite of 2.3).

To sum up, the clay minerals consist of platy crystallites, the lattice structure of which is made up of layers of alumina and silica linked by oxygen bridges and which contain hydroxyl groups, some of which are active in base exchange and are at least in part located on the surface of the particles.

Plasticity.

While the plasticity of clay-water mixtures is one of the most important properties of the clay, it is so complex in nature that no one has been able to develop for it a satisfactory quantitative formulation or method of measurement. On the other hand, the more important fundamental factors affecting it are clear.

If one mix a dry clay with increasing amounts of water, a certain minimum is necessary to give coherence to the mixture.[*] If the water be further increased, the mass is capable of plastic flow, with a yield value that decreases progressively. A point is finally reached at which the yield value disappears.[†] The interval between the two critical water contents is the plastic range of the clay. Indeed, the difference in these has been suggested as a coefficient of plasticity, but unfortunately does not always parallel plasticity in its significance to the potter. Meanwhile, if clay in the plastic range of water content be deformed, excessive distortion or distortion rate can cause loss of cohesion, with a resultant disintegration that makes it unworkable for the potter. Finally, even though workable, it is necessary from the practical point of view that the molded clay should possess a considerable degree of mechanical strength. A plastic clay is one possessing capacity for plastic flow over wide

---

[*] With less water, the clay crumbles. The quantity of water required for cohesion depends on the conditions: thus, under high pressure a mass, otherwise crumbly, will stick together satisfactorily.

[†] Unfortunately, this point is difficult to measure accurately.

ranges of water content,* combined with ability to withstand high distortions, and yet at the same time mechanically strong. Furthermore, there is always an optimum water content for distortability combined with mechanical strength. The difficulty in combining all these characteristics in a single measurement or numerical coefficient is obvious. Plasticity clearly represents a compromise of properties more or less independent of each other.

It is probably still true that, despite the personal factor, the most practical method of evaluating plasticity for the potter is by the feel of the wet clay to the hand. Mobilities and yield values as functions of water content can be measured by the conventional methods (p. 139), but throw no light on coherence. One can roll the plastic mass into the form of small cylinders or wires, and determine the diameter below which disintegration results. One can also measure torsional distortion of a cylinder of the plastic mass as a function of the shear and in specific cases each of these quantitative results may have value.
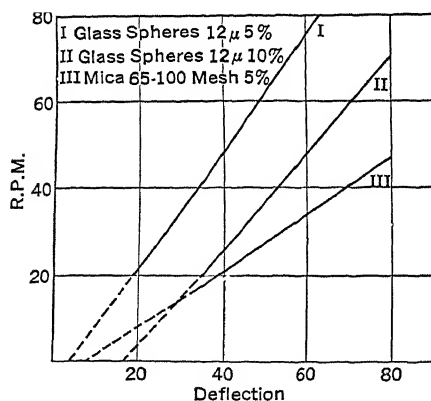
FIG. 2. Consistency Curves for Mica and Glass Spheres in an Acetylene Tetrachloride-Monobrombenzene Mixture of Equal Density.†

The mechanism of plasticization of clays is in dispute, but it seems probable that the major factor is mechanical interference of platy particles dispersed in the water. It has been demonstrated that plastic flow can be induced in liquid suspensions of asymmetric particles, whether needle-shaped or platy, or even mechanical agglomerations of spherical particles (Fig. 2). Asymmetric particles can develop mechanical structure in the mass at extraordinarily low volumetric concentrations, particu-

---

* The significant point is probably the ratio of water to clay.
† BROUGHTON, G., and WINDEBANK, C. S., *Ind. Eng. Chem.* **30**. 407 (1938).

larly if the ratio of the linear dimensions be high.   Particle size is of major importance; the finer the particles the lower is the volumetric concentration necessary for a given plasticity (Fig. 3).

It seems certain that the ratio of interparticle attractions to the distorting forces likely to act on a single particle increases as size decreases.   This is equivalent to saying that forces holding the ultimate particle in place decrease less rapidly with size than do those inducing movement.

FIG. 3. Influence of Particle Size upon Flow Characteristics.*

Dispersion of the clays is essential to plasticity.   It is connected with affinity of the liquid for the clay surface; thus clays resist dispersion in all organic liquids of low polarity.   Dispersion is promoted by wet grinding and by moderate rise in temperature.   It is one of the major results of ageing and weathering.   There is every reason to believe that, when adequately dispersed, the surface of each ultimate particle is highly hydrated, even where agglomeration is considerable.   However, auxiliary factors frequently have large influence.   Thus, a relatively dilute electrodialyzed bentonite suspension having an undetectable yield point will set to a strong gel on addition of KOH or KCl.   The effect of barium hydroxide is far greater, but that of caustic soda less † (see Figs. 13 and 14, Chapter X). These compounds, if added in high concentration, flocculate the bentonite, causing it to settle and making it incapable of gelation if kept in suspension at the original concentration.   Indeed, it is apparently a general rule that addition in small amount of a solute which in relatively high concentration will result in

---

* BROUGHTON, G., and HAND, R. S., Am. Inst. Mining Met. Eng., *Tech. Publication* No. 1002 (1938), and *Petroleum Technology*, Nov., 1938.

† The value of much of the data on additions to natural clay suspensions is limited because of lack of information regarding the metals already combined with the clay.

flocculation tends to develop yield value. However, the phenomena are far more complicated. A suspension of a natural clay can also undergo a very great reduction in yield point on addition of traces of alkali salts of certain high molecular weight acids, *e.g.*, the sodium tannates, silicates, and hexametaphosphates. In addition to possible flocculating or deflocculating effects (*i.e.*, influence on particle adhesion) it may be that specific solutes profoundly modify interparticle friction. There is some parallelism between the base-exchange capacity of clays and plasticity (Table II, in which the clays are arranged in order of increasing plasticity).

TABLE II.   BASE-EXCHANGE CAPACITY AND PLASTICITY

Clay	Base-Exchange Capacity, in milli-equivalents per 100 gms.
Kaolin	5
Ball clay	13
Fire clay	17.5
Fuller's earth	27
Bentonite	90

Because air dispersed through the clay-water mixture can greatly decrease plasticity, de-aeration is being adopted in making refractories, bricks, and pottery.

Some natural clays are found in beds in which they have apparently been naturally dehydrated to a considerable degree. These have capacity for rehydration and plasticization, but frequently the rate at which this occurs is extremely slow. The clays are mixed with water and stored in thin banks, the process being allowed to go on sometimes for as long as one to two years. When openly exposed to the weather, in addition to rehydration, rain can wash out soluble salts, including a portion of the iron content, while frost and cold help to crack up large agglomerates.

Fusibility of Clay.

If one heats a pure clay mineral it loses water, as indicated in Fig. 1, breaking down to alumina and silica. On further heating, minor changes occur, detectable by thermal effects, but ultimately liquefaction sets in, according to the curves of

Fig. 4. Inspection of this figure shows that by control of temperature alone one can secure any desired degree of lique-faction of the mass, except at the monovariant points. If liquefaction be kept within moderate limits there remains suffi-cient solid, skeletal material so that the hot mass will hold its
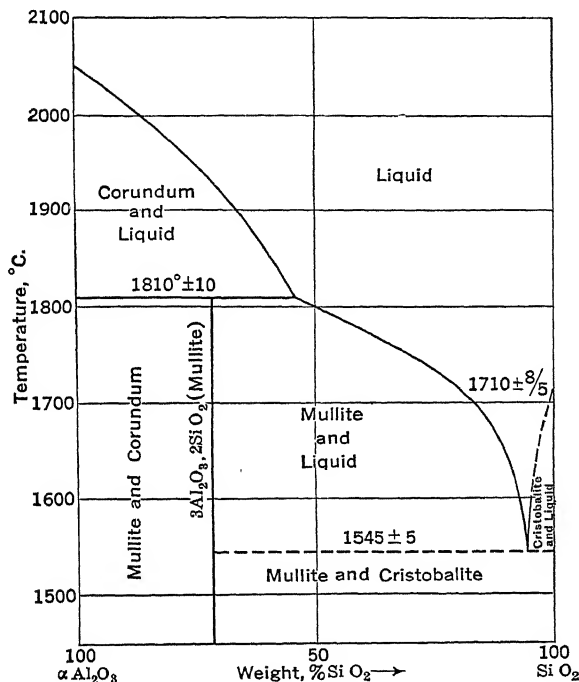
FIG. 4. The System Al_2O_3–SiO_2.*

shape, and on cooling one will have a structure bonded together by the vitrified portion of it. This behavior is the basis of all ceramic manufacture.

While the lowest temperature at which pure alumina-silica mixes can liquefy is 1545° C., the impurities present in natural clays always have a lower melting point or induce it by eutectic action. Furthermore, the molten constituents, *e.g.*, feldspar, are likely to be powerful fluxes for both silica and alumina (cf. glass). Because of this influence on fusibility it is scarcely too much to say that, from the point of view of the potter, the im-

* BOWEN, N. L., and GRIEG, J. W., *J. Am. Ceram. Soc.* 7, 238 (1924).

purities in a clay are likely to be its most important constituents. There are two major classifications of ceramic products depending upon the degree of vitrification attained: porous and vitreous wares.

Manufacturing Technique

Despite the diversity of ceramic products, the steps in manufacture are standardized to a high degree. The essential operations are plasticizing the mix, forming the article, drying, and burning; the two last sometimes are combined. Auxiliary steps frequently employed are the blending of various clays, etc., in the mix, glazing the articles, decoration, and the like; these, too, are often combined with others.

Clays are crushed and screened to free them from pebbles and gross impurities, and weathered when necessary. Fine clays are ground with water in wet pans or ball mills, dispersion being aided where necessary by addition agents, such as alkali (p. 242). Gritty impurities, such as silica, can now be removed by settling in a classifier or otherwise. The clay thus purified can be coagulated with acid and collected on filter presses.

For manufacture of high-grade products it is nearly always necessary to blend various raw clays, as well as to incorporate other mineral constituents required in the mix, e.g., feldspar, flint, etc. This is essential to secure the requisite structure and degree of vitrification in the ultimate product, rarely achievable by the use of a single natural clay. So important is this that, once the potter has found a satisfactory blend, it is rarely modified, other than by minor changes in proportions required by fluctuation in quality of the raw materials. However, in blending, it is equally important to control the plasticity of the mix from the points of view of both workability and shrinkage on drying, which more or less parallels plasticity. Since the strains induced in drying are likely to cause cracking, the potter always compromises between high plasticity for moldability and low for drying. Plasticity can be controlled by choice of the clays employed in the mix, such as the plastic, "fat" ball clays or the contrasting "lean" kaolins, but this is clearly not independent of burning characteristics. However, by introducing ground, burnt material of the same composition, so-called "grog,"

plasticity can be reduced independently to any desired degree. Breakage and wastage around the plant is usually employed in this way, although it is obviously desirable to keep the amount to a minimum. Occasionally additional grog must be burnt. In shaping ceramic articles, clay is manipulated either in the form of a stiff paste or as a thin, fluid suspension, or "slip." The paste can be prepared in a pug mill, a trough equipped with a powerful central screw or shaft and blades which knead the mix with water, propelling it slowly to the discharge end, at which it is ejected by extrusion. After pugging, the clay is often aged for a matter of months in a damp atmosphere at 25 to 30° C., to secure further increase in plasticity.* For slips, the clay, *e.g.*, from the filter press, is stirred or "blunged" with water to the desired concentration, the flow characteristics being controlled for maximum concentration compatible with satisfactory "pourability," by addition agents, such as sodium silicate (p. 248).

The clay paste may be shaped by hand on the potter's wheel, a revolving table. This operation, old in tradition and requiring considerable skill, has been largely modified and replaced in modern plants. Thus, "jigging," an operation particularly suited for the mass production of simple shapes, *e.g.*, plates, saucers, etc., involves forming the paste in a rotating mold, the stock conforming to the mold under a combination of centrifugal action and mechanical pressure. The attention of the operator is thus required only for a single surface. Another important technique is casting, employing clay slip. This is poured into plaster of paris molds, so porous that the water penetrates them rapidly, producing on the surface a filter cake of clay. When this cake has built up to the desired thickness, the excess slip is poured out. Drying of the mass shrinks the clay sufficiently to allow the articles to be removed readily from the mold and leaves the latter dry for re-use.† Certain articles, such as bricks and tiles, may be formed by extrusion under pressure. The pugged clay is fed to an auger, which in effect is a pug mill

---

* Living organisms undoubtedly play an important part in this operation, although the mechanism is far from clear.

† Electrophoretic deposition of the clay from a slip on the mold as an electrode has been suggested.

capable of exerting extremely high extrusion pressures. The mass flows as a continuous slab through an oil-lubricated die, wire cutters being used for division into correct lengths. Where precision of shape is essential the bricks are pressed in molds, frequently after some drying or "tempering." A technique,
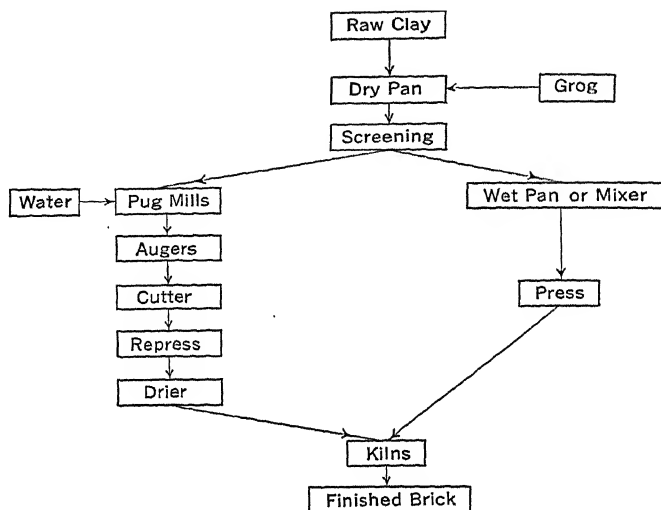
FIG. 5.   Flow Sheet of Brick Production.

growing in importance, is so-called dry-pressing, in which the mix moistened with insufficient water to form a plastic mass (5 to 10%) is pressed in die-molds at a sufficient pressure (500 to 1500 lbs. per sq. in.) to secure cohesion. This reduces shrinkage on drying to a point such that the stock can frequently go directly to the kiln.

Drying.

Since moisture can escape from the clay ware only by evaporation from the surface and can travel from the interior only under the driving impetus of concentration gradient,* complete elimination of shrinkage strains on drying is impossible. They can, however, be reduced to extremely low values by taking

---

* Movement of liquid through the stock is predominantly a question of capillary action, but study of the mechanism makes it clear that in drying ceramic ware a concentration gradient is in general necessary,

sufficient time on drying, under temperature and humidity control that will eliminate moisture irregularities at the surface. This control, together with thermal efficiency, is best achieved by the use of counter-current driers, preferably of the tunnel type. The more plastic the mix and complicated the shape, the more careful must be the drying operation.* Occasionally, drying can be carried out in the kiln itself, utilizing the low temperature sensible heat in the waste gases where the kiln is countercurrent in action.

Burning.

Because clay is a poor conductor of heat, rapid heating, impossible without a high temperature gradient through the stock, must be avoided at all costs, since the resulting internal strains lead to warping, cracking, and breakage. In the burning operation, the initial temperature rise expels residual moisture from the stock and finally the chemically combined water; this period is known as "water smoking." Care must be taken that gases from the hotter parts of the kiln are not cooled below their dew point by contact with cold stock elsewhere. This results in moisture deposition on the cold ware, which may weaken or blemish the articles, as by scum formation on the surface due to reaction of sulfur oxides from the combustion gases with alkalies or alkaline earths in the clay. Water smoking usually requires at least fifteen to twenty-four hours, but is substantially complete below 800° C., when the ware usually contains less than 1% moisture.

The next stage in firing is at a temperature level sufficient for oxidation by excess air of impurities in the articles which might otherwise cause serious trouble. Almost all clays contain carbonaceous matter, which, if oxidation be not completed while the mass is still relatively porous, can reduce iron oxide when present, giving black cores. Blebs, small bubbles due to gas evolution, may also form. It is desirable, therefore, to hold the ware at a temperature level of 800 to 900° C. for a time sufficient to allow diffusion of the oxygen necessary for oxidation.

---

* Not only does shrinkage in drying parallel plasticity of the wet paste, but the percentage of voids in the dried mass does likewise; these may amount to 60% of the total volume. High voids obviously result in further severe shrinkage on burning.

In the final stages of firing, a portion of the mix is fused to serve on cooling as a bond for the whole, leaving, however, an adequate skeleton of unfused material as a framework to retain shape (see p. 327). There is a minimum of fusion which must occur to give the necessary bond for the stock in question, and
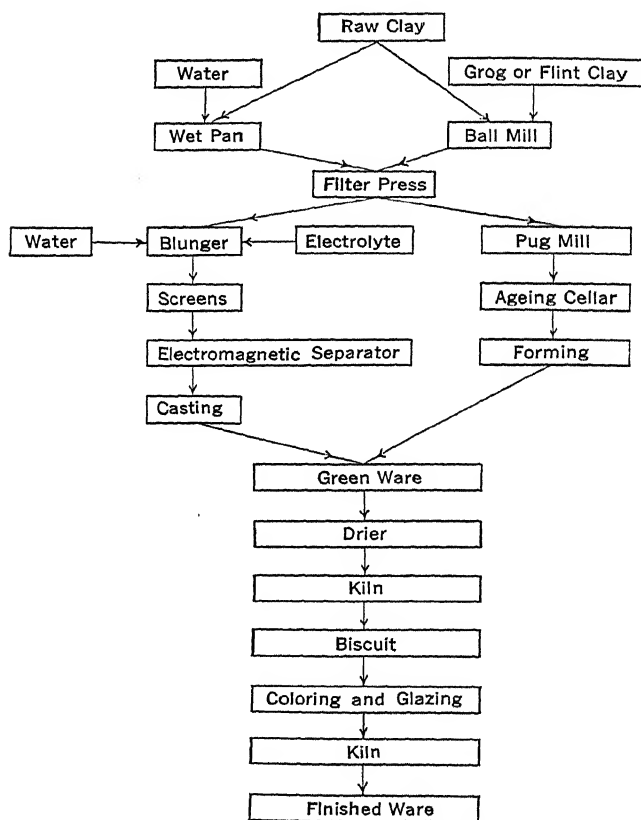
Fig. 6. Flow Sheet of Ceramic Ware Production.

a maximum beyond which shrinkage or even softening may develop, or insufficient porosity result. For a satisfactory product every portion of the ware must be burnt between these tolerance limits. These, in turn, usually correspond to reasonably specific lower and upper temperatures, the interval being the allowable firing range. This obviously requires careful

temperature control.* Common brick and stoneware may be burnt at temperatures as low as 850° C., but the temperatures may range as high as 1100 to 1250°. Ordinary pottery requires 1050 to 1250° and porcelain 1300 to 1500°. Firebrick may be burnt as low as 1200° but the highest quality fire-clay products go up to 1700°. Where the pottery will be injured by direct contact with combustion gases or unusual uniformity of temperature is desired, it is packed before introduction into the kiln in saggers, rough clay boxes re-used repeatedly until breakage occurs.

Kilns.

Rough ware such as brick can be burnt in the simplest of structures, sometimes built of unburnt brick itself. On the other hand, careful, uniform burning requires provision for meeting the conditions just outlined. An important type of batch kiln, shown diagrammatically in Fig. 7, consists of a circular firebrick chamber covered with a refractory, dome-shaped roof. Fuel is burnt in the fireboxes at the base of the kiln, the combustion gases flowing upward around its whole inner periphery or through flues in the walls, hitting the roof and being deflected downward through the ware, finally leaving through ports in the floor which lead into horizontal flues going to the stack. The kiln itself is to a large degree heat jacketed

FIG. 7. Batch Kiln.

by the hot combustion gases around it. The most important technique of temperature control is regulation of excess air. This is initially very large to avoid overheating the "green" ware. The excess is progressively reduced until burning is com-
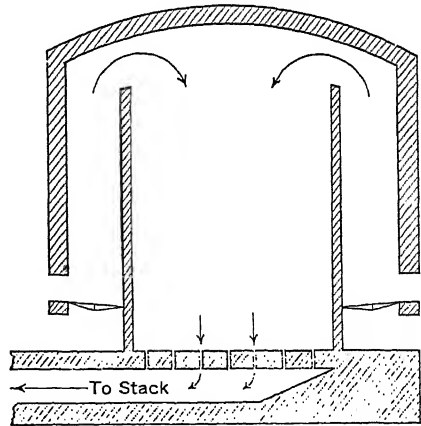
* This is especially important in vitreous ware, where it is necessary to obtain low porosity by melting a high percentage of the mix. It is obvious that a relatively small increase can result in softening and collapse.

pleted. The combustion gases of necessity go to the stack at the temperature at which they leave the ware. This, together with the large excess air used, particularly in the earlier stages of burning, results in extremely low thermal efficiency. However, such kilns are flexible and capable of thoroughly satisfactory control.* None the less, their low heat efficiency has necessitated the development of the continuous type of kiln.

The principle of the continuous kiln is most readily illustrated by the tunnel type. In this the stock, supported on cars, is fed relatively continuously through the long tunnel. The running gear of the cars is protected from the high temperature of the firing zone, e.g., by a refractory skirt dipping in a sand seal. Counter to the charge flow the combustion gases. Air enters at the stock discharge end and flows over and around the cars moving in the opposite direction. About the middle of the tunnel the fuel is introduced, preferably as gas, and in any event carefully distributed throughout the firing zone so as to minimize local overheating. The combustion gases preheat the incoming green ware. Because the heat effects of the changes occurring in the final burning of the ware are limited and the sensible heat required to raise its temperature is theoretically recoverable, the minimum quantity of heat necessary for burning is extraordinarily small. Indeed, it is so small that the air necessary for theoretical combustion of this minimum quantity of fuel is completely inadequate to absorb the sensible heat in the burnt stock. Consequently, satisfactory heat utilization requires the flow through the kiln of quantities of air far in excess of those required for combustion of the fuel employed. This, however, has the great practical advantage of furnishing a large quantity of gas for distribution over the stock at all points, making it possible to secure uniformity of heating at each stage in burning, the necessity for which has already been emphasized.

A semicontinuous type is the older ring kiln, in which a number of batch kilns are built together, so that they can be operated on a cycle. The waste gases from the hottest of the kilns in which the combustion occurs are used to preheat the

* Thus, in this type the downward flow of the gases, which are falling in temperature and hence increasing in density, distributes them evenly and minimizes localized temperature fluctuations.

green ware in the cooler units, while the incoming air for combustion is preheated by passage over the cooling finished ware. The ring kiln has the advantage of operability at very high temperature levels, at which the operation of cars becomes difficult.

An important technique for controlling the burning of ceramic ware is the use of Segar cones, small narrow-based pyramids made up of carefully controlled ceramic mixes of well defined softening temperatures. These can be located at various points inside the kiln where they can be observed, their collapse occurring when that point in the kiln has reached the corresponding temperature. They have a certain time lag, but are cheap enough to be distributed freely around the kiln and are free from the high temperature deterioration suffered by many high temperature measuring instruments.

Glazing.

Because all true ceramic ware is porous to a certain degree, complete impermeability can be secured only by coating it with a thin film of glass, an operation known as glazing. The glaze must have a fusion temperature lower than the softening point of the ware, must bond satisfactorily with it, and possess a coefficient of expansion over the whole temperature range up to its own softening point sufficiently close to that of the ware to avoid cracking or flaking. The former defect, "crazing," appears to be due to excessive, and the latter, "shivering," to insufficient contraction of the glaze. The glaze also should be resistant to water and other destructive agencies, uniform in composition, and of satisfactory appearance—clear, opaque, or colored, as may be required.

The simplest glaze to apply, often used for stoneware, is a "salt-glaze," obtained by throwing salt into the hot kiln. The salt vaporizes sufficiently to react with the surface of the ware, forming mixed sodium aluminum silicates which lower the melting point sufficiently to form a thin, transparent, surface coat. Because control of such an operation is difficult, high-grade ware is always glazed by application of a suitable mixture to the surface of the ware before introduction into the kiln. To secure

uniformity of composition and application, the glazing mixture is semifused or "fritted" in a separate operation and converted to a fine powder by grinding.   Suspended in water the powder is applied to the surface of the ware by brushing, spraying, dipping, or similar techniques.   The glaze is often applied to the green ware after drying, but this makes independent control of the burning of the body of the ware and of the glaze impossible. Hence, high-grade pottery is first burnt unglazed and the frit applied to the burnt stock or "biscuit," which is then reburnt in the "glost" kiln.   This makes it possible to secure a somewhat coarse structure between the body and the glaze, limiting contact and improving the stability of the bond between the two.

Glazes are complex in composition, frequently containing B_2O_3 (see p. 304), oxides of lead, tin, and zinc, and the like.   The flow characteristics of their aqueous suspensions may be controlled by addition of electrolytes.   Iridescence can be produced by exposing the hot ware to vapors of stannic or ferric chloride.

Decoration is usually secured in connection with glazing, the colors being applied either in or beneath the glaze.   In the former case they may function as pigments or as color dissolved in the glass of the coating itself (p. 299).   While pottery decoration offers opportunity for artistic work of high order, mass production requires mechanical reproduction of design by the use of stencils, transference from decalcomania paper (soaped paper to which the design is applied by an etched plate or rolls), and the like.

Refractories.

An important branch of the ceramic industry is the manufacture of refractories, *i.e.*, of ceramic shapes for use in furnace construction, capable of withstanding physical strain and chemical action at high temperatures.   Refractories must consist of skeletal material of high melting point, preferably chemically resistant, cemented together with a minimum of fusible bonding material.   Because impurities usually tend to lower the melting point, highly refractory materials always consist of relatively pure substances.   The melting points of some of the more important refractory bases are given in Table III.   In many

instances, however, the melting point of a refractory is not a true criterion of its value. Magnesite brick, for example, has a melting point of over 2000° C., but has little load-carrying capacity above about 1500° C., while, on the other hand, mullite brick will carry a load to within fifty degrees of its melting point, 1810° C.

TABLE III.   MELTING POINTS OF REFRACTORY BASES

Base	Formula	Chemical Character	Melting Point, in °C.
Fire clay		Neutral	Up to 1750
Silica	SiO_2	Acidic	1710
Chromite	$FeO.Cr_2O_3$	Neutral	About 2000
Magnesia	MgO	Basic	2800
Mullite	$3 Al_2O_3.2 SiO_2$	Neutral	1810
Carborundum	SiC	Neutral	Dissociates 2200
Alumina	Al_2O_3	Neutral	2020
Zirconia	ZrO_2	Neutral	2700

Firebricks are the refractories produced in largest amount. Figure 4 is the phase rule equilibrium diagram for the system Al_2O_3–SiO_2.   Above 5% Al_2O_3, the greater the amount of alumina present, the higher the fusion range of the refractory: furthermore, only one compound, mullite, is formed.   Clearly, mixes richer in alumina than 72% will not begin to melt till above 1810° C., provided equilibrium, with its attendant mullite formation, is attained.   Consequently, mixtures of synthetic alumina with natural clays high in alumina are being used increasingly for refractories.   To mold refractories, the original mix must be plastic, and for this purpose refractory, plastic clays are usually employed.   This clay itself serves as the ultimate bonding material for the skeletal materials of high melting points, the refractory being burnt at a temperature sufficient to melt the clay and to cause it to etch and bond together the refractory body. In order to get mechanical strength and resistance to chemical erosion at high temperatures, maximum density is desirable in the ultimate particle of the refractory base itself.   This can sometimes be secured by repeated high temperature sintering and regrinding of the skeletal materials.   Density of the mix can be increased by using skeletal particles of properly assorted size (as is done in concrete mixing), filling the voids between the

larger particles with material of smaller dimensions. The whole is finally bonded together wet with a minimum amount of plastic clay, and molded or dry-pressed.

Some types of refractory may be bonded chemically. Silica brick is often bonded with lime, which on firing forms fused calcium silicate to act as a bond between the silica particles. Magnesite brick may also be chemically bonded with a paste consisting of chrome ore, clay, and sodium bisulfate.* Carborundum bonds at a high temperature under conditions that oxidize the surface of the grains to silica, which, softening, cements together the whole mass, forming so-called autogenously bonded carborundum bricks. Bonding may also be secured by adding sufficient alumina to combine with the silica formed by oxidation to synthesize a mullite cement.

Refractories should be burnt for some time at temperatures above the highest to which they are to be submitted in use Fluxing action of the bond on the skeletal material keeps up at high temperatures for a long time, thus increasing the amount of fused material with corresponding shrinkage of the refractory and increase in density. It is obvious that this process should be completed so far as practicable before the refractory is built into the ultimate structure in which it is to be employed.

BIBLIOGRAPHY

CURTIS, E. DEF., *Pottery, Its Craftsmanship and Its Appreciation*, Harper (1940).

FORSYTH, G. M., *Art and Craft of the Potter*, Chapman and Hall, Ltd., London (1934).

GRANGER, A. A., *Céramique Industrielle*, Gauthier-Villars et Cie., Paris (1929).

JENKINS, R. H., *Practical Pottery for Craftsmen and Students*, Bruce (1941).

NORTON, F. H., *Refractories*, McGraw-Hill (1931).

PARMALEE, C. W., *Clays*, Edwards Bros., Ann Arbor, Mich. (1936).

RIES, H. R., *Clays*, Wiley (1927).

DE SAGER, W. H., *Making Pottery*, Studio Publications, New York (1934).

SEARLE, A. B., *Chemistry and Physics of Clays*, D. Van Nostrand (1924).

SEARLE, A. B., *Encyclopedia of Ceramic Industries*, E. Berm, Ltd., London (1929).

WILSON, H., *Clay Technology*, McGraw-Hill (1927).

* Prepared in this way it can be used without preliminary firing.

Chapter XX

Synthetic Resins and Plastics

Before the development of the methods of synthesis of high molecular weight polymers described in Chapter VII, the manipulation as plastics of natural materials of this type almost always involved a certain degree of disintegration of the molecular structure (as in the solution of cellulose and the milling of rubber), and involved only to a limited extent the building up of structure in the finished product (*e.g.*, the vulcanization of rubber and the drying of unsaturated oils). Since the development of adequate methods of polymerization, industry has been to an increasing degree in a position to produce materials of almost any required physical characteristics, possessing chemical resistivity in high degree, and yet to manufacture them under conditions such that they could be manipulated as plastics at intermediate stages in their synthesis. Of these materials the most important developed to date are the synthetic resins.

It is important to re-emphasize the fact that the combination of mechanical strength with high physical and chemical inertness and resistivity seems to require high molecular weight. This can be secured by building up one of two types of polymer, either linear or three-dimensional in structure (see p. 157), the mechanism of polymerization giving, in the final stages, far greater molecular weight for a given amount of chemical action in the three-dimensional type. Furthermore, while linear molecules may possess great mechanical strength, particularly if combed out more or less parallel to each other, they are capable of relative motion and slip. Hence, they tend to be flexible and, because of their limited molecular weight, soluble and fusible; three-dimensional polymerized molecules, on the other hand, because of their interlocked structure and larger molecular

weight are strong but tend to be brittle, extremely insoluble, and difficult to soften by high temperature short of decomposition. In consequence, the characteristics and industrial uses of these two types of polymers differ.   In both types polymerization is induced, at least to a considerable degree, prior to final manipulation and shaping of the ultimate article.   If plasticization by solvents (treated elsewhere) be excluded, both types of polymers are usually manipulated while softened by heat.   The linear types are polymerized to the final stage before manipulation. The requisite plasticity for shaping is then secured by heat, which, however, produces no permanent change in properties. In the three-dimensional plastics, initial polymerization is carried as far as practicable, but final cross-linkage is secured by heat after shaping, giving a product no longer softened by heat, owing to the irreversibility of the chemical changes which occur in ultimate set.   The first type is therefore classified as *thermoplastic* and the second as *thermosetting*.   The product of the initial stages of polymerization of the thermosetting resins may be either a liquid or a solid, but in the case of thermoplastic materials it must be solid.   Particularly where the solids are to be employed as bonding agents for fillers, it may be desirable to prepare them as powders for convenience in mixing; the mixture is then used in pressure molding.

Initial polymerization is a chemical reaction, the conditions for which depend upon the individual case and are described under the specific product.   All the processes are polymerizations or condensations, with a velocity of reaction influenced greatly by temperature, catalysts, and, occasionally, light. Catalysts frequently affect not only the velocity but also the nature of the reaction and consequently the character of the product (p. 478).   Negative catalysts or inhibitors of the polymerization desired are also known and must be carefully excluded.   True polymerization has the obvious advantage that it produces no by-product.   Hence, commercial processes involving condensation are preferably conducted so that as large a quantity as possible of the water or other product eliminated in the initial stages is removed prior to final set.*   This is often

* It is desirable to control condensation so as to minimize water formation during final set.

accomplished by evacuation, allowing time for escape of water from the material.

Mechanical Processing of Resins.

Perhaps the most important function of the synthetic resins is as bonding agents, either for comminuted fillers, wood flour, asbestos, mica, pigments, etc., or for fabrics (paper, cloth, etc.). In the former case the finished object is usually shaped by molding, casting, or extrusion; in the latter, the material in sheet form is impregnated with the resins by dipping. Except where solvents are used, plasticization is effected by heat, ultimate deplasticization being secured either by cooling or by continued thermal polymerization.

In the molding of thermoplastic materials, the stock must be held under pressure after attainment of the required plasticization temperature for a time sufficient to enable it to conform fully to the mold shape. If, at this molding temperature, plasticity be sufficient to cause deformation after removal from the mold, the object must be cooled while still under pressure. In molding thermosetting materials, the temperature must be sufficiently high to secure adequate plasticity for molding without reaching a polymerization rate that will stiffen the article excessively before conformation to the mold is completed, *i.e.*, there must be sufficient setting time. The operation therefore represents a balance between the temperature coefficients of plasticity and the polymerization rate of the material in question. Indeed, this factor determines the allowable degree of preliminary polymerization of the material prior to final molding and cure. Once set has occurred, the shorter the final curing time, the better.

Molding is used for articles such as electrical parts, telephone receivers, bric-a-brac, etc. For mass production the time factor is important. For small articles a cure cycle as short as a minute may be sufficient. The molds, usually of hardened steel and sometimes chromium plated, are mounted in hydraulic or mechanical presses, capable of exerting from 1 to 4 tons per sq. in. For limited production, removable hand molds are employed; but on a large-scale, semi-automatic presses are used,

in which the molds are built in and do not have to be withdrawn by the operator. In the former the platens of the press are generally heated by steam; in the latter type steam and water can be alternately circulated through the channels. Ejection of the molded pieces is preferably automatic. The two main types are compression and injection molding (Figs. 1 and 2). Compression molding may be "overflow" or "positive." In the former as pressure is applied a slight excess of charge is squeezed out, producing a fin or flash where the mold sections meet. This overflow material is subsequently removed by grinding or filing. In positive molding the plunger fits tight, allowing no escape of excess material, and thus requiring careful adjustment of the charge. In injection molding the charge is plasticized by heat in a chamber external to the mold and forced into the latter, as by the plunger of Fig. 2.† The technique lends itself to mass production and is advantageous for thermoplastics.
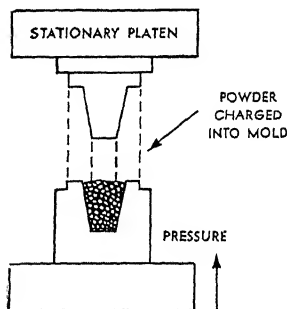
FIG. 1. Compression Molding.*
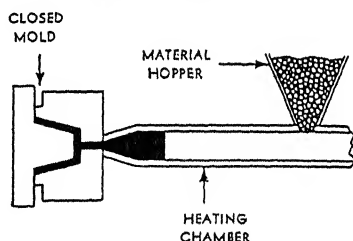
FIG. 2. Injection Molding.*

Some resins may be formed by casting the still-liquid polymer into molds without pressure, polymerization being completed by slow baking. The product is often machined to give the desired shape.

Laminated materials are prepared by impregnating sheets of cloth or paper with partially polymerized material; application of heat to the piled sheets finally produces a dense, strong, completely bonded product. The resin is usually prepared in the form of a varnish with 40 or 50% solvent, which is added to the reaction kettle at completion of the first stage in the synthesis

* *Bakelite Molded Plastics*, p. 7, Bakelite Corporation (1940).
† Figure 2 illustrates repeated charging of the mold by a single charge of the heating chamber. Where the heating chamber is recharged for each filling of the mold, the operation is usually called "transfer molding."

of the resin.* The fabric is impregnated and the solvent driven off by drying at a temperature low enough to avoid set-up.† Cut and piled in multiple sheets, the impregnated fabric is introduced into a press, protected by sheets or pans of stainless steel at each end of the pile. Commercial presses may accommodate sheets as large as 4 by 8 ft., using pressures up to 1500 lbs. per sq. in.‡ Material can be fabricated which is characterized by extraordinary impact strength and wear resistance, suitable, *e.g.*, for gears.§

High-grade laminated products demand careful control of the plasticity of the resins used for impregnation, to avoid squeeze-out under pressure in the early stages of cure without sacrifice of impregnation. The method of plasticity control may be determinations of either resin solubility, percentage of volatile constituents, or viscosity. As in molded materials, the resin content of the product varies between wide limits, 10 to 60%, depending on the characteristics required. The resin percentage is controlled by thinning the varnish with solvent, or by passing the sheet through squeeze rolls before drying.

Phenol-Formaldehyde Resins.

Phenol, cresols, and other phenolic bodies when heated with formaldehyde in presence of catalysts give thermosetting resins, representing the earliest industrially important synthetic development (1909). When heating is continued for sufficient time and an excess of formaldehyde is present, insoluble resins decomposing without fusion at about 300° C. are obtained (see p. 317). The reaction is complicated by the fact that at least two distinct paths, summarized in Fig. 3, may be followed to the final infusible product. In the first, called somewhat illogically the one-step process, slightly more than one mol of formaldehyde is allowed to react in presence of a basic catalyst with one mol of the phenol. Reaction proceeds readily and may be stopped

---

* For this case basic catalysts are usually used, rather than the acid catalysts common in preparing molding compounds.

  † The drying oven may be horizontal or vertical, the latter being less convenient in operation but conserving floor space by using several passes of the sheet.

  ‡ With the aid of steel molds and mandrels, shaped laminated objects can be made.

  § Thermoplastics, *e.g.*, cellulose acetate in acetone solution, are sometimes used for impregnating paper or cloth.

at any point by lowering the temperature. Short reaction times produce a product which is solid but soluble and fusible, known as a "resole" or A-stage resin. Its average molecular weight is

$$* C_6H_5 . OH + H_2C:O$$

$$HO . C_6H_4 . CH_2OH$$
Hydroxybenzyl Alcohol

Excess CH$_2$O
Basic Catalyst

Excess C$_6$H$_5$. OH
Acid Catalyst

CH$_2$OH
$$-[C_6H_2(OH)-CH_2-]_n$$
Resole (A Stage)

OH
$$-CH_2-[C_6H_3-CH_2-]_n C_6H_3-$$
OH
Novolak

Heat

Hexamethylene Tetramine Heat

Infusible Resin (C Stage)
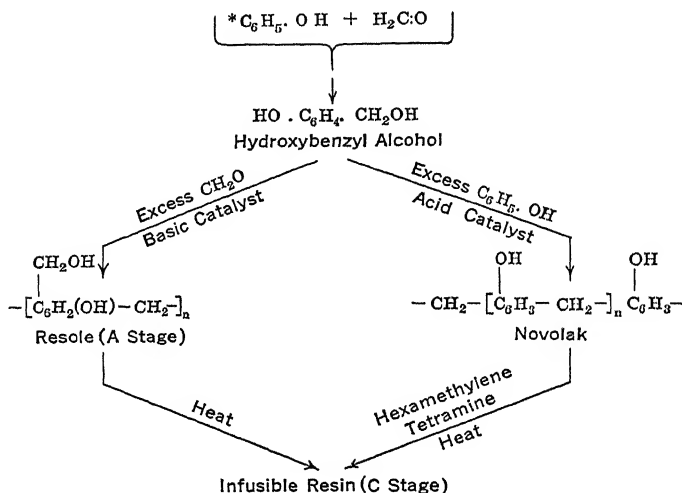
*Substituted phenols may also be used, provided reactive hydrogens (ortho and para) are left available on the ring.

FIG. 3. Phenol-Formaldehyde Condensation.

small, of the order of 300 to 400. On reheating, reaction recommences, finally yielding an infusible (C-stage) resin,* stable against further reaction below decomposition temperatures.

In the so-called two-step process, 1 mol of the phenol is heated with 0.7 to 0.8 mol formaldehyde in the presence of an acid catalyst, giving a fusible resin, known commercially as a "novolak," of average molecular weight 700 to 1000, somewhat higher than that of the resole. This, on heating with more formaldehyde, conveniently supplied in powder form as hexamethylene tetramine, yields insoluble, infusible resins as final products, apparently identical with the C-stage resin from the one-step process.

Although the mechanisms are not fully understood, it is believed that they are fundamentally similar in both processes. Taking a phenol-formaldehyde mixture as an example, hydroxy-

* Intermediate products, which may for example be insoluble but considerably swollen by solvents, are sometimes called B-stage resins.

benzyl alcohols are first formed, substitution occurring only at positions ortho and para to the hydroxyl group,

$$C_6H_5.OH + CH_2O \rightarrow HO.C_6H_4.CH_2OH.$$

Alkalies catalyze this reaction, but are believed to have little effect upon the succeeding reaction by which a linear, soluble polymer of limited length is formed. The first step would be the following:
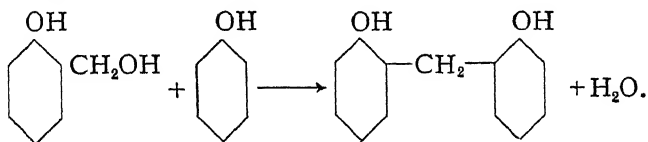
Subject to the limitations stated, this can again react with more formaldehyde and phenol to give a structure of the type
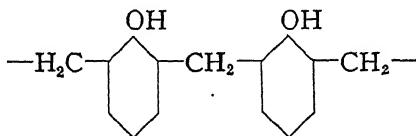
However, para-substitution also occurs, correspondingly complicating the product. Granting that alkalies minimize chain development relative to alcohol formation, while acids catalyze both reactions, promoting activity of the nuclei on the ends of the chains, alkalies should encourage formation of compounds such as:
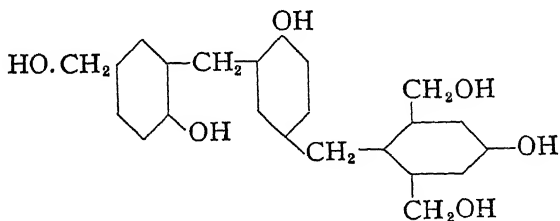
in contrast with the following type of structure for acid catalysis
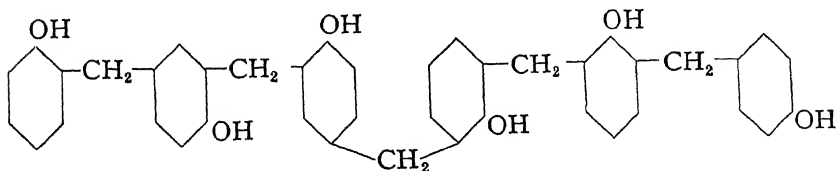
In either process, cross-linkage finally occurs to give a three-dimensional polymer:

$$
\begin{array}{c}
\text{(three-dimensional cross-linked phenol-formaldehyde polymer structure showing linked phenol rings bearing OH groups connected by } CH_2 \text{ bridges)}
\end{array}
$$
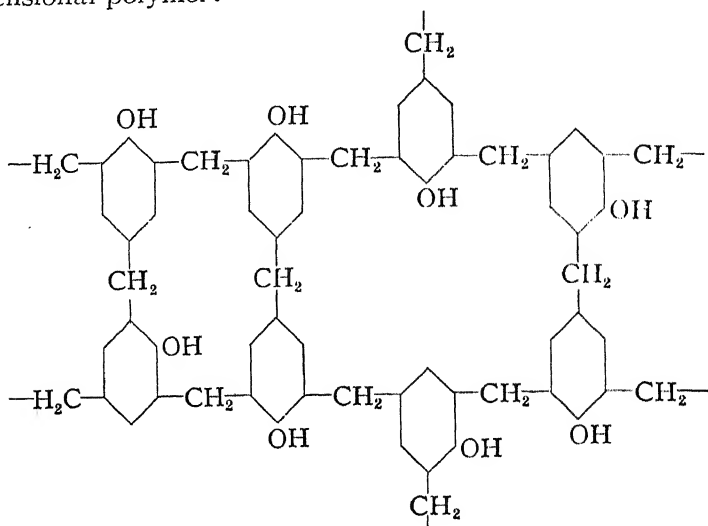
In general, 1 mol of phenol, 1.5 or 0.75 mols of aqueous formaldehyde (dependent upon whether a basic or acid catalyst is used), and catalyst are placed in a closed, steam-jacketed kettle and heated for several hours. Water is finally eliminated by heating under vacuum and the molten resinoid removed from the kettle. Usually acid catalysis is used for quick-curing molding compositions, subsequently incorporating hexamethylene tetramine in the molding powder, while resins for varnishes (see below) and laminated products are made with alkaline catalysts. Resin and filler, with pigments or dyes to give color, are ground together and run between hot rolls to insure thorough impregnation of filler with resin. While all fillers reduce shrinkage, each type has special advantages; thus, fiber gives better impact resistance, asbestos better heat resistance. Wood meal is frequently used. The resulting sheets may be used as such, or when ground to a powder constitute the usual molding material. By casting at relatively low temperatures without fillers, incompletely dehydrated, light-colored, cast products are obtained. These may be either opaque or clear, depending on elimination of water, and lend themselves when dyed to beautiful

color effects.   As a class, these resins have poor color stability to both light and heat * but possess good electrical insulating qualities, undergo little shrinkage on setting, and hence are widely used for insulation, brackets, handles, etc.   With finely divided asbestos as filler the phenol-formaldehyde resins give a mechanically strong, heat-resistant material, often used in chemical construction work, *e.g.*, acid-resistant tanks.

Urea-Aldehyde Resins.

Heating a mixture of urea (or thiourea) and formaldehyde in the presence of hexamethylene tetramine produces a water-insoluble body, which may be dehydrated and molded under heat and pressure to yield a hard, infusible, and highly transparent product.   Like phenol-formaldehyde resins, it is a condensation polymer.

$$
\begin{array}{ccc}
\quad NH_2 & NH_2 & NH.CH_2OH \\
\diagup & \diagup & \diagup \\
CO \quad + CH_2O \to CO & & \to CO \\
\diagdown & \diagdown & \diagdown \\
\quad NH_2 & NH.CH_2OH & NH.CH_2OH
\end{array}
$$
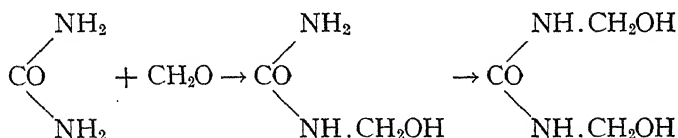
These particular compounds, mono- and dimethanolurea, have been isolated.   Further condensation with reaction of hydrogens of the imino groups may lead to compounds such as the following:

$$
\begin{array}{cccc}
& CH_2 & & CH_2 \\
& | & & | \\
-CH_2-N-CH_2-N & NH-CH_2-N \\
& | & | & | & | \\
& CO & CO & CO & CO \\
& | & | & | & | \\
& NH-CH_2-N-CH_2-N-CH_2-N \\
& & & & | \\
& & & & CH_2 \\
& & & & |
\end{array}
$$
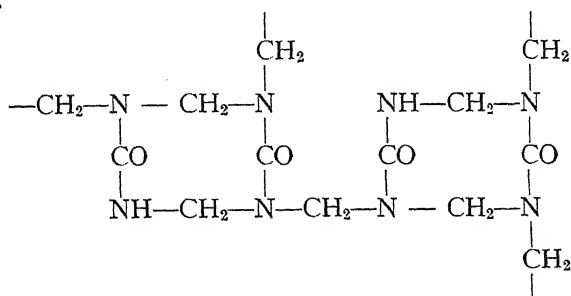
The reaction is very sensitive to hydrogen-ion concentration, temperature, and proportions of the reactants, a mixture of not over 3 mols of formaldehyde to 1 of urea apparently giving best results.   Nickel or chrome-nickel kettles are used since iron

* They can be used in light, decorative colors if protected from exposure to direct sunlight and heat.

discolors the resin. Water elimination in urea-formaldehyde resins gives trouble. It is important to get rid of all water formed in the preliminary polymerization, but even so it is difficult to avoid a certain degree of opacity in thick articles cured at high temperatures (see, however, p. 470). These resins are valuable, because they are odorless, stronger than phenolic resins, and permit a wide range of transparent or translucent dyed colors, fast to light.* They have found application, therefore, in the fabrication of tableware, radio cabinets, buttons, trays, and ornaments; a more recent development is in the production of creaseless fabrics (p. 490). They are, however, liable to develop internal strains especially in thick shapes,† but filler, e.g., bleached sulfite pulp, reduces the difficulty. Moreover, they become opaque slowly above 80° C.

Vinyl Resins.

Long-chain, addition polymers can be prepared from certain compounds containing the vinyl group ($-CH=CH_2$). They therefore behave as thermoplastics.

$$n \ CHR=CH_2 \rightarrow -CH-CH_2-CH-CH_2-CH-CH_2- $$
$$\qquad\qquad\qquad\quad | \qquad\qquad | \qquad\qquad | $$
$$\qquad\qquad\qquad\quad R \qquad\qquad R \qquad\qquad R $$
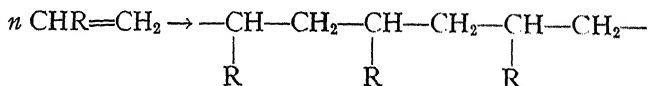
For satisfactory polymerization the radical R should be of a negative character. The reactions have received considerable study because of increasing technical importance (Chapter VII). The catalysts used greatly influence the character of the final product. Styrene (p. 147), for example, when allowed to polymerize at room temperature with no catalyst, gives a fibrous, white solid, almost insoluble in ether, of molecular weight approximately 500,000. Strong sulfuric acid or antimony pentachloride as catalyst at the same temperature gives saturated, soluble polymers of molecular weights about 1000 and 5000, respectively. Stannic chloride leads to a soluble polymer, still unsaturated, of molecular weight 3000 to 7000. Peroxides and oxygen are milder in action, leading to polymers of molecular weight 20,000 and over. Temperature likewise exerts an im-

* Laminated sheets in direct sunlight discolor somewhat.
† Perhaps due in part to the problem of water elimination.

portant influence both on rate of reaction and nature of the polymer formed. Inert solvents slow down polymerization rate and also decrease the average molecular weight of the polymer formed. Dilution of partially polymerized styrene with mono-styrene does not appear to affect the mean molecular weight of the final polymer produced. Similar relationships seem to hold for other vinyl polymers.

Styrene, the source of the important polystyrene resins, is made by direct thermal dehydrogenation of ethylbenzene,

$$C_6H_5.CH_2.CH_3 \rightarrow C_6H_5.CH{=}CH_2 + H_2,$$

or by chlorination of ethylbenzene, followed by elimination of hydrogen chloride,

$$C_6H_5.C_2H_5 \rightarrow C_6H_5.C_2H_4Cl \rightarrow C_6H_5.CH{=}CH_2 + HCl.$$
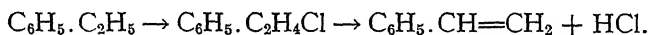
Polymers of molecular weights of 20,000 and over are of industrial utility. They are characterized by high transparency and strength with excellent dielectric qualities, even better than those of shellac, making them useful for high frequency insulation. They can be worked by a number of methods. Plasticization by solution followed by evaporation leads to ribbons and thin films. Many articles can be molded, products of high molecular weight requiring higher molding temperatures. Thus, while 120° C. is a suitable temperature for a resin of mean molecular weight 60,000, 155° C. is needed for 150,000. Polystyrene can also be extruded and machined.

The vinyl resins proper are formed by polymerization of esters of vinyl alcohol, catalysts such as aluminum chloride, boron trifluoride, acetyl and benzoyl peroxides being used. Vinyl acetate is obtained as a low-boiling liquid by reacting acetylene with acetic acid in the presence of mercuric sulfate or other catalyst:

$$CH_3COOH + CH{\equiv}CH \rightarrow CH_3.COO.CH{=}CH_2.$$
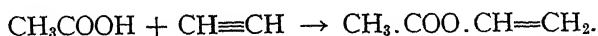
It is polymerized by refluxing in benzene solution containing a trace of acetyl peroxide to give a fusible, somewhat oil-soluble, transparent, thermoplastic resin. Molecular weight and, consequently, physical properties are controlled by the proportion of

catalyst added.   Vinyl chloride also gives a resin, which is hard, brittle, and much less thermoplastic.   Co-polymerization of vinyl chloride and acetate yields a mixed resin, known under the trade name, vinylite, which combines the more desirable properties of both.   Furthermore, the properties can be varied for specific purposes by varying the relative proportions of chloride and acetate.   The vinyl ester resins are odorless, resistant to moisture, dilute acids, and alkalies, and undergo little shrinkage on setting.   Properly prepared and plasticized, they retain toughness at low temperatures and are in consequence especially valuable in safety glass inter-layers.

Acrylic esters, *e.g.*, methyl acrylate, can be synthesized by hydrolysis of acrylnitrile in the presence of the alcohol.   The nitrile can be formed by dehydration of the cyanhydrin of either acetaldehyde or ethylene (the latter formed from alkali cyanide and ethylene chlorhydrin or hydrocyanic acid and ethylene oxide).   Methyl acrylate polymerizes to a stable, valuable resin. Methyl methacrylate, synthesized by similar reactions from acetone cyanhydrin, forms a thermoplastic polymer, not only stable, strong, and tough, but unique in freedom from color, in transparency and color stability, so that it promises to replace glass for many optical uses.

Alkyd Resins.

The alkyd resins, known under the trade names glyptal and rezyl, are of the condensation type, formed by esterification of polybasic acids with polyhydric alcohols by heat (pp. 140–146). Glycol and phthalic anhydride * are limited to the formation of linear polymers only:

$$-OOC \quad COO.CH_2.CH_2.OOC \quad COO.CH_2.CH_2.OOC \quad COO-$$
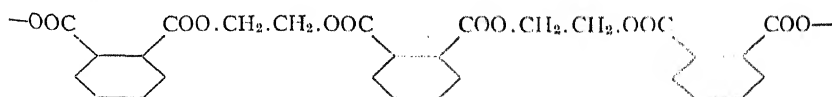
However, as discussed on p. 155, glycerol and phthalic anhydride can give a polydimensional resin of which the following ultimate structure is presumably typical (glycerol and phthalic anhydride

---

* The industrial availability of phthalic anhydride (*e.g.*, from oxidation of naphthalene) is far greater than that of suitable aliphatic dibasic acids.   This has been an important factor in promoting its use for resin syntheses.

residues being represented by G and P respectively):

$$-\text{G}-\text{P}-\text{G}-\text{P}-\overset{|}{\text{G}}-\text{P}-\overset{|}{\text{G}}-\text{P}-\text{G}-$$
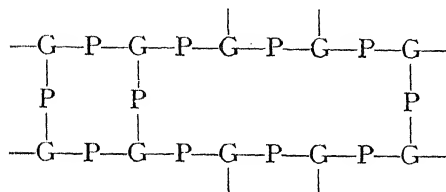
Unmodified resins produced by heating together 1.5 mols of phthalic anhydride and 1 mol of glycerol are hard and infusible and but little used. Commercially, the resin is modified by the introduction of monobasic acids, which function as softeners. Such acids, by reacting with the hydroxyl groups of the glycerol, obviously limit the scope of polymerization, but the products are important in varnishes (p. 317), although little used as molding compositions. The properties vary with the nature of the modifying agent used. Baking finishes, exceptionally resistant to heat discoloration, moisture, and light, can also be obtained by using these resins. The ester is applied in the soluble A-stage, polymerization to the infusible C-stage resin being caused by the heat treatment. Such finishes have good adhesion.

The variations in the important physical properties of the more widely used synthetic resins are surprisingly small. The specific gravities range from a little over unity for polystyrene to about 1.8 for some of the phenolic resins. The refractive indices are of the order of magnitude of 1.45 to 1.55. The tensile strengths are from 3000 to 10,000 lbs. per sq. in.; where plasticization is essential, the tensile strength, of course, tends to fall. The modulus of elasticity ranges from 100,000 lbs. per sq. in. to about 500,000, although that of the urea-formaldehyde resins may go to 1,500,000. The molding temperature range is of necessity narrow; if it is below about 140° F. the thermoplastic resins are too soft at normal temperatures and the thermosetting materials are likely to give trouble by hardening before they are ready to mold. At temperatures above 350° F., thermal decomposition is likely to ensue. Good resins should be water-resistant, but the cellulose esters will sometimes absorb as much as 8% in twenty-four hours at room temperature, although the

corresponding absorption of many of the other resins may be as low as a few hundredths of 1%.

Hitherto the most important single source of the basic raw materials for the manufacture of the synthetic resins, excluding cellulose derivatives (discussed in Chapter XVI), has been the by-products of distillation of coal. Phenols, cresols, and xylols are used directly in the bakelite type of resins. Naphthalene is oxidized in the vapor phase, with vanadium pentoxide as catalyst, to phthalic anhydride for use in the alkyds. Benzene is a source of synthetic phenol, in addition to that present in the coal tar as such; it can be reacted with ethylene to form styrene: it is oxidized catalytically in the vapor phase (*e.g.*, with vanadium pentoxide) to maleic anhydride, which in turn can be hydrogenated to succinic acid, both for use in the alkyds. On the other hand, all these coal tar products must be supplemented by other materials. A second important source is acetylene, obtained in turn from calcium carbide. Acetylene is converted by acetic and hydrochloric acids into vinyl acetate and chloride, respectively, the monomers of the vinyl resins. It is converted through vinyl acetylene to chloroprene (p. 447); it is hydrated catalytically to acetaldehyde, which in turn yields butadiene by dehydration. By oxidation of the acetaldehyde to acetic acid and conversion of the latter to acetone, it serves as a source of the methacrylates through the cyanhydrin reaction. The third important source of raw materials is the high-pressure syntheses of ammonia and methyl alcohol from hydrogen, the latter reacting in the first case with atmospheric nitrogen and in the second with carbon monoxide, the ammonia of the former reaction being used for conversion of carbon dioxide to urea and the methyl alcohol of the second operation for oxidation to formaldehyde. A source, so far almost untapped, which promises to be of outstanding importance in the field, is the by-products of cracking petroleum. By proper choice of the stock treated and of the cracking conditions, it is possible to secure good yields of such important hydrocarbons as ethylene, isobutylene, butadiene, and even acetylene. While these are produced directly as individual components of complex mixtures, requiring rather complicated systems for segregation and purification, the fact that other

valuable products of the pyrolysis can help carry the cost of manufacture makes this type of synthesis unusually promising. Indeed, progress has already been made in modifying the pyrolysis of oil in the manufacture of illuminating gas so as to produce important quantities of valuable unsaturated hydrocarbons, such as styrene.

It is worthy of note that the only commercial, basic raw materials needed for the manufacture of the whole group of substances just referred to are fuel in some suitable form (serving as a source not only of the essential hydrocarbons but also of the energy required) and, in the case of the acetylene derivatives, limestone.

Casein Plastics.

Proteins react with formaldehyde, possibly according to the equation,

$$R.NH_2 + H.CHO \rightarrow R.N{=}CH_2 + H_2O,$$
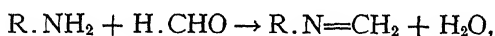
to give hard, stable products. This is the basis of casein plastics used for such articles as buttons and buckles. The casein is ground, mixed with plasticizers and water to a uniform dough, and pressed into sheets or extruded under pressure through a die. Paper pulp may also be incorporated as a filler. The shapes thus formed are allowed to stand in formaldehyde solution for periods up to several months until completely hard.

CELLULOSE ESTER AND ETHER PLASTICS (See Chapter XVI)

Plasticizers.

The importance of plasticizers in the compounding of polymers and resins can scarcely be overemphasized. They act predominantly as solvents * for the polymer, functioning as intermolecular lubricants and thereby greatly reducing brittleness. Because they reduce tensile strength (an effect which may be obscured by the reduction in brittleness), they must be used in limited amount. To be permanent, they must be nonvolatile and chemically stable. They must have little or no tendency to crystallize. They should not be susceptible to the action of solvents with which the finished product will come in

---

* Substances of this type which are not solvents for the polymer but which are imbibed by it, swell it and presumably dissolve in it, are sometimes called softeners.

contact.* These requirements limit the number of available materials. Many of them are high boiling esters, such as triphenyl phosphate, the dialkyl phthallates and sebacates, triglycerol hexoate and octoate and methyl stearate.† Historically, camphor as used in celluloid was one of the first, and in many ways is unique in action.

The use of synthetic resins is developing rapidly. The advantages of laminated gears have already been mentioned. Synthetic resin bearings are coming into use, even for heavy duty operation. The combination of strength and acid-resistance makes the resins suitable for articles such as rayon centrifugal pots (p. 364). Electrical insulation was among the earliest applications but its importance is still increasing. They also make possible decorative effects combined with durability and strength hitherto unknown. Future expansion of the industry will be limited only by the possibility of reducing costs.

BIBLIOGRAPHY

ELLIS, CARLETON, *The Chemistry of Synthetic Resins*, Reinhold (1935).

GOGGIN, W. C., and LOWRY, R. D., Vinylidene Chloride Polymers, *Ind. Eng. Chem.*, **34**, 327 (1942).

HOUWINK, R., *Physikalische Eigenschaften und Feinbau von Natur- und Kunstharzen*, Akad. verlag. m.b.h., Leipzig (1934).

Modern Plastics, Monthly magazine, Breskin and Charlton Publishing Corp., New York (October issues, 1936–1939, contain very helpful bibliographies).

MORRELL, R. S., ed., *Synthetic Resins and Allied Plastics*, H. Milford, Oxford University Press, London (1937).

RAHM, L. F., *Plastic Molding*, McGraw-Hill (1933).

ROWELL, H. W., *The Technology of Plastics*, Plastics Press, Ltd., London (1936).

SIMONDS, H. R., *Industrial Plastics*, Pitman (1939).

SUTERMEISTER, E., and BROWN, F. L., *Casein and Its Industrial Applications*, Am. Chem. Soc. Monograph, Reinhold (1939).

U. S. TARIFF COMMISSION, *Synthetic Resins*, Report No. 131 (1938).

High Polymers: A series of monographs on the Chemistry, Physics, and Technology of High Polymeric Substances is being published by the Interscience Publishing Co., Inc., New York.

* In specific cases the plasticizer can reduce the sensitiveness of the resin to solvent action.

† *Ind. Eng. Chem.*, **34**, 68 (1942).

Chapter XXI

Textile Fibers

The textile fibers are not manipulated as plastics (although some are thus synthesized), but are spun into yarn and woven into fabrics. However, even in these mechanical operations successful utilization requires the knowledge of their structure and properties. The origin and classification of the more important fibers will be clear from Table I.

COTTON

As the cotton plant nears maturity, the seed pods or bolls break open, and to a considerable degree the individual hairs dry and collapse to produce long, narrow bands, which take on somewhat of a longitudinal corkscrew curl, in addition to the transverse curl around the axis of the fiber induced by the flattening on drying [Fig. 1 (*a*) and (*b*)]. The extent to which each fiber is made up of interwoven individual fibrillae is brought out beautifully in Fig. 2, made under polarized light. The fibers vary in length from about 1.3 to 3 cm. for ordinary staple to 4 to 5 cms. for Egyptian or Sea Island cotton. Spinning is assisted by the natural convolutions of the fiber, which vary from about 60 turns per cm. in Indian cotton to 120 in the long staple types.

The raw fibers are separated from the seeds in the gin, caught up by rotating, saw-tooth discs, from which they are subsequently removed by compressed air or by a cylindrical brush rotating more rapidly than the disc. The short immature hairs, called linters, are removed in a second, more drastic ginning operation. Linters are worthless as a textile material, but due to their availability and comparative purity are used extensively, after a preliminary purification, as chemical raw material for the

manufacture of celluloid, smokeless powder, lacquer base, rayon, and other cellulose derivatives.

Cotton Bleaching.

Although it is one of the purest forms of natural cellulose, a typical raw cotton contains only about 85% of the carbohydrate, together with a fraction of 1% of oils and waxes, about 5% proteins, pectoses, and coloring matter, a trace of mineral salts, and about 8% water. For most textile uses these materials must be removed, but this is rarely done before spinning and often not until after weaving.* To avoid loss in strength (tendering), the purification must be conducted to minimize degradation by either hydrolysis or oxidation (p. 154). In general, it involves two distinct steps: a high temperature treatment with dilute alkali, the kier boil, followed by a hypochlorite bleach

(a)

(b)
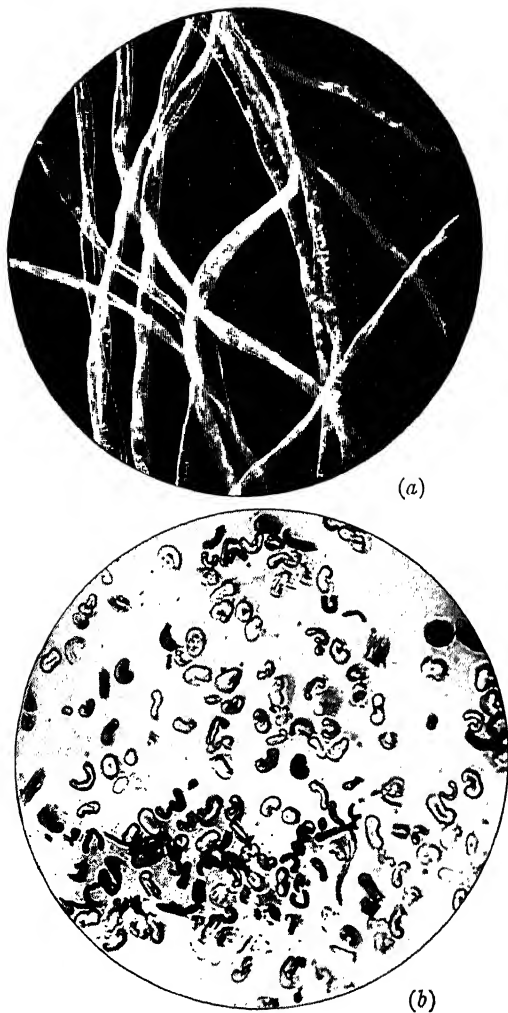
Fig. 1. Photomicrographs of Raw Cotton ($\times$ 200).†
(a) Longitudinal section
(b) Cross section

* Thus, the natural waxes are valuable as fiber lubricants during spinning.
† Courtesy of Prof. E. R. Schwarz, Textile Laboratory, Massachusetts Institute of Technology.

(chemicking).   The former hydrolyzes and emulsifies fats and waxes and dissolves pectic and starchy materials; the latter destroys or dissolves coloring matter.*

FIG. 2.   Fibrillar Structure of Cotton Fiber in Polarized Light (× 800).†

A typical operation for bleaching gray cloth is the following. The chemical operations as such are usually preceded by singeing, passing the dry cloth at high speed over a Π-shaped metal surface heated to a dull red heat to burn off protruding fibers and nap.‡   The cloth is now roped by passing through a porcelain ring or "pot-eye" and wetted out by drawing under rolls through a 2% solution of caustic soda.   It is then plated down into a kier, a steel cylinder 10 to 12 ft. in height and 6 ft. in

AIR VALVE
STEAM VALVE
MAN HOLE
BAFFLE PLATE
FALSE BOTTOM
DRAIN-OFF PIPE
STEAM INJECTOR
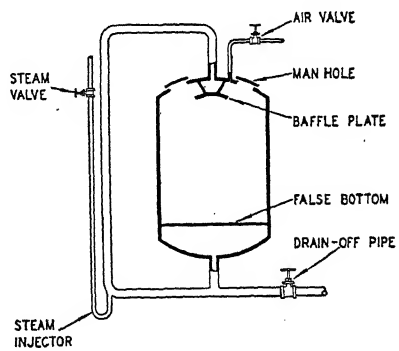
* The color bodies are sometimes merely bleached and not removed.
† Courtesy of Prof. E. R. Schwarz, Textile Laboratory, Massachusetts Institute of Technology.
‡ It also tends to break down bits of boll, leaf, etc., which may remain in the fiber, so that they can be dusted out.

diameter, provided with a false bottom (Fig. 3). After the kier is filled, the manholes are carefully closed and liquor is circulated through the stock from bottom to top by a steam injector or otherwise. This "bowking" is continued for several hours at approximately 30 lbs. pressure. Absence of air to avoid degradation by oxy-cellulose formation is essential. Perfect circulation of the lye through the stock is also necessary. At the end of the operation the lye is drained off and wash water introduced.*

The goods are roped out of the kier, sometimes washed, but in any event impregnated with a solution of bleaching powder of 0.5 to 1° Tw.† The impregnated cloth is laid down in piling pits and allowed to stand for several hours, carefully protected against drying out. The cloth is then again picked up in rope form, carried through washers, "soured" in 1 to 2° Tw. hydrochloric or sulfuric acid, again washed, squeezed, opened, and dried. In connection with the final washing an antichlor, e.g., sodium thiosulfate, may be used to insure complete elimination of chlorine. Despite the mechanical differences, the bleaching of cotton goods closely parallels that of wood pulp (p. 342). The oxidizable impurities in cotton are smaller in quantity, but the fiber seems more sensitive to bleaching conditions. In particular, degradation of cotton is serious where the bleaching is conducted below a pH of about 9. Evidently because of these low hydrogen-ion concentrations longer times are required for the action of hypochlorite on cotton.

Mercerization.

Mercerization is a treatment which results in greatly increased luster, a soft, silky feel, and an increased power of response to mechanical finishing processes. If a single fiber of cotton be treated in 14 to 28% sodium hydroxide solution it swells, becomes gelatinous, and, when the caustic soda is washed out, is found to have shrunk about 20% in length. Examined under the microscope, it is thick and transparent, nearly cy-

---

* It was formerly common practice to conduct the alkali boil with lime, this being usually followed by a secondary boil with sodium carbonate solution.

† In the textile trade specific gravities of solutions are frequently reported as degrees Twaddell, each degree corresponding to an excess of 0.005 specific gravity above unity.

TABLE I. THE TEXTILE FIBERS *

Textile Fibers	Fiber	Usual Length of Fiber	Diameter, in mm.	Kgs. per sq. mm. Dry	Kgs. per sq. mm. Wet	Grms. per Denier Dry	Grms. per Denier Wet	Extension Dry %	Extension Wet %	Standard Regain, % (60–70% R.H.)	Principal Uses	Origin
Cellulose — Natural	Cotton	½″–2″	0.010 –0.023	43 (21–80)[a]	47 (24–82)	3.2	3.5	8.5	9.0	8.5	All classes of cloth and union goods; Raw material for many chemical processes	U. S. A., India, Egypt, Brazil
	Flax		0.013 –0.025	84	88	6.3	6.6	1.8	2.2	12.0	Linen	Ireland, Russia, Belgium, etc.
	Hemp		0.015 –0.05	90	—	6.8	—	1.7	—	12.0	Cordage, twines, canvas, etc.	Russia, France, Japan, India, etc.
	Ramie		0.02 –0.075	93	108	6.6	7.8	2.3	2.3	—	Surgical materials, gas mantles	India, Japan, China, U. S. A.
	Jute (contains lignin)		0.0015–0.02	44	—	3.4	—	1.3	—	13.75	Sacking, oil-cloth, rugs, twine, etc.	India
	Nitro	As desired	As desired	13	—	—	—				Obsolete	
Cellulose — Synthetic	Viscose			{ 29	16 (9–20)	2.3	1.3	16	25	11.0	All types of textiles	
				61	53	4.5	3.9	9	8.9			
	Acetate			19 (16–20)	11.5	1.6	1.0	26	29	9.0		
	(Cuprammonium)			28 (23–33)	18 (15–21)	2.0	1.3	18	24	11.0		
Protein — Natural	Wool	1″–15″	0.05 –0.086	17 (13–22)	15 (11–17)	1.4	1.2	37	47	14–19	Worsteds, woolens, and union goods	
	Silk	Cocoon 3000–4000 yds.; 600–1200 yds. reeled	0.018	55	50	4.5	4.0	15	30	11	Clothing of all kinds	Japan, Italy, China, France, India
Protein — Synthetic — Lanital {casein, wool}		As desired	As desired	11	5.6	0.92	0.47	9	100	—	Generally used mixed with other fibers because of low strength	
Silicate — Natural — Asbestos (3MgO·2SiO₂·2H₂O)		As desired									Packing, fireproof materials, insulating fabrics	Canada, Siberia, Corsica, Australia, etc.
Silicate — Synthetic — Glass Fiber		As desired	As desired	50 (1–8 denier)	47			14	30		Filter cloths, fancy effects, etc.	

Note: In the Silicate natural row the formula reads $3MgO \cdot 2SiO_2 \cdot 2H_2O$.

[a] Figures in brackets indicate the extreme values observed for any given kind of fiber.

* Data of Schmidhauser, O., *Melliand Textilber.* **17**, 905 (1936).

lindrical in section, and has completely lost its former character-
istic twist. If a yarn or fabric be submitted to this alkali
treatment, conducted under tension, shrinkage of the fibers can
be reduced, and the mass acquires a lustrous sheen; merceriza-
tion has occurred. The more regular surface of the fiber thus
treated gives it a superficial resemblance to silk.

Cellulose has a marked affinity for caustic soda (p. 360), but
does not form a definite chemical compound. After washing
out the caustic soda the X-ray diffraction pattern of the cellulose
is identical with that obtained from regenerated celluloses, e.g.,
viscose. The pattern indicates that the cellobiose chains have
been turned through a small angle, giving it an inner micellar
surface considerably greater than that of ordinary cellulose.
This would account for the increased activity of mercerized
cotton towards reagents of all kinds, notably dyestuffs and water
vapor. That degradation does not occur is shown by the fact
that copper number and viscosity in cuprammonium solution
remain substantially unchanged after mercerization, while
tensile strength may actually increase.

It is immaterial whether the cotton is prevented from shrink-
ing throughout the process or stretched to its original length
while being washed. Consequently, an arrangement of the
type shown in Fig. 4 is used industrially. The goods are im-

pregnated with about 25% sodium hydroxide solution by means
of a mangle E, then pass over a series of rolls F (to secure even
penetration) to the moving clips of the stenter frame G. These
securely grasp the fabric, preventing it from shrinking, while
sprinklers H, above and below the cloth, wash out the caustic
soda. Suction boxes (cf. p. 355) have been used for the more
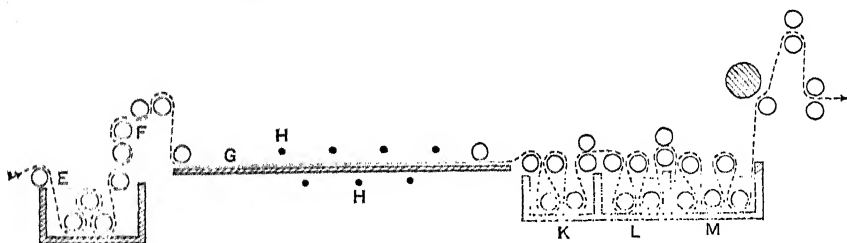
efficient extraction of caustic soda. The cloth finally passes to a washing tank K, is soured by dilute acid in L, and washed in M. Speeds of upwards of 60 yds. per minute with 90 to 95% soda recovery can be obtained, countercurrent washing being employed.

Finishing.

Cotton fabrics are frequently sized or loaded. The most important sizing material is starch, introduced by impregnation followed by drying and calendering. Stiffeners or fillers, *e.g.*, barium or calcium sulfate or china clay and the like are frequently incorporated with the starch. The size cements the fibers together, thereby increasing tensile strength, gives a glossy surface finish, and increases resistance to soiling. The loading materials fill the spaces between the fibers, interfering with flexing and thereby stiffening the fabric.

During bleaching, both warp and weft suffer distortion which tends to remain after fabrication. Because the major distortion is lengthwise elongation due to handling in rope form, it has long been the practice to correct this by tentering,* the edges of the wet cloth being gripped in an endless chain of clips which hold the cloth under sidewise tension while it is being dried. How-
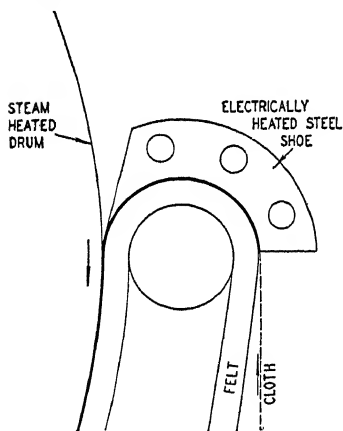
FIG. 5. Technique of Sanforizing.†

ever, the cloth is still left under a strain which causes shrinkage on washing, and correction is desirable. In sanforizing the cloth is dried while under a certain degree of lengthwise compression. The moist cloth is fed to the machine of Fig. 5 between tenters set, not to induce sidewise stretch, but rather to control the degree of sidewise contraction resulting from the moderate lengthwise tension on the strip at this point. This gives the weft threads a definite crimp. The cloth is now laid

* Also called stentering.
† RAMSEY, C. H., *Trans. Amer. Soc. Mech. Eng.* 59, 335 (1937).

down on the elongated convex surface of a felt traveling over a roll of high curvature, and is held by shoes with sufficient pressure to prevent slip. Curvature is reversed by bringing the sheets against the surface of a drum. The resulting contraction of the felt surface compresses and crimps the warp of the cloth with no further effect on the weft. The degree of foreshortening of the warp is controllable by adjusting the relative speed of the felt and the second drum. This makes it possible to set the cloth with any desired amount of crimp in both the warp and weft threads. While on subsequent washing true shrinkage occurs to a certain degree, this is compensated by an accompanying straightening out of the threads so that net shrinkage is kept to a minimum.

The final operation of cloth finishing always involves calendering between heavy rolls, plain or embossed, cold or steam-heated, depending on the specific case.

LINEN

The growing flax plant contains just under the bark a layer of long cellulose filaments, each made up of small, individual cells about 0.00125 to 0.025 cm. in diameter and 0.6 to 4 cms. long, cemented together by an intercellular, pectous material. After removal of the seeds, branches, etc. ("rippling"), the fibers are loosened from the bark or cortex and the hard center by "retting," a selective putrefaction induced by soaking in water. This breaks up woody impurities, which can be removed by beating, leaving the bast fibers undamaged.* Products spun or woven from these are known as linen.

Flax fibers are straight and polygonal in cross section [Fig. 6 (a) and (b)]. Because of their length they can be spun into far finer threads than cotton. The cellulose crystallites appear to be better oriented in linen than in cotton, its tensile strength being greater and extensibility less (Table I). In agreement with this view, linen shows no improvement in properties upon mercerization. Although flax is more difficult to bleach than cotton, the same methods may be employed. Linen is some-

---

* If the flax plant be allowed to mature and go to seed (as in linseed oil production) the plant becomes woody and so resistant to retting that isolation of the fibers becomes impracticable.

times bleached by "crofting"—spreading out for a period of weeks or months in the sun—supposedly giving a superior product. Bleaching is avoided when maximum strength is required.

RAMIE, HEMP, AND JUTE

Like flax, hemp, jute, and ramie are bast fibers, the first two being likewise extracted by retting. By scraping the bark off the ramie plant, ramie fiber can be obtained as brown ribbons. Boiling with caustic soda extracts the brown coloring matter, leaving white fibers of high, almost silky luster. Ramie possesses durability but little flexibility, tending to split when bent suddenly. Like the other bast fibers it has little extensibility.

Hemp is coarser than linen. Its excellent strength is seriously affected by bleaching, which, therefore, is seldom carried out.

Jute differs from the other bast fibers in the extreme shortness of its individual cells
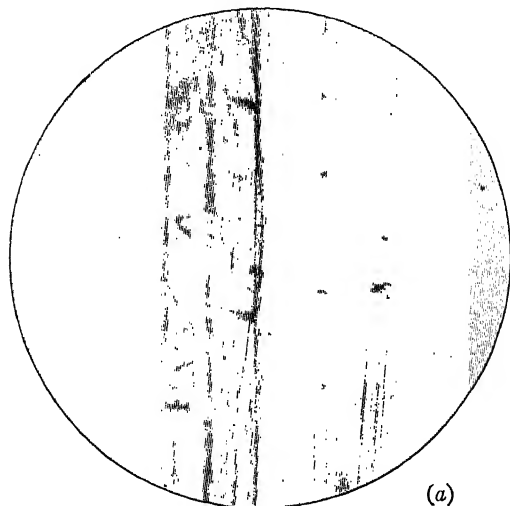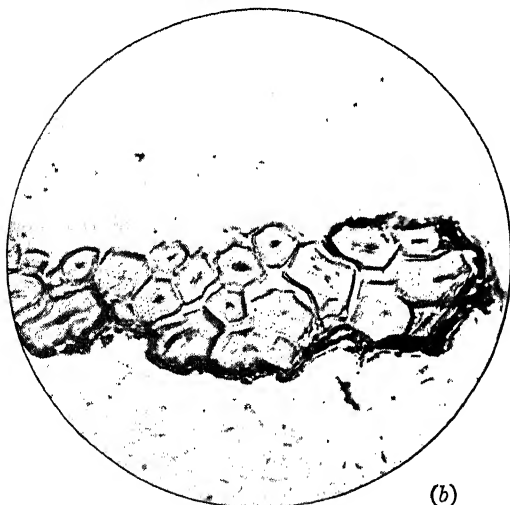
FIG. 6. Photomicrographs of Courtrai Flax ($\times$ 300).*
(a) Longitudinal section
(b) Cross section

* Courtesy of Prof. E. R. Schwarz, Textile Laboratory, Massachusetts Institute of Technology.

(1/16 to 3/16 in.) and its high proportion of lignin, because of which it is extremely difficult to bleach. On account of its harsh feel, coarseness, and, particularly in the presence of water, lack of durability, it is used for the manufacture of articles where strength is not a primary consideration.

WOOL

Wool fibers grow in the hair follicles of the skins of sheep, etc. (cf. Fig. 4, p. 394), and are associated with considerable impurity, about 8 to 18% wool fat or grease (lanolin) and 15% dried sweat ("suint"), derived from the sebaceous and sweat glands, respectively. In addition, considerable quantities of dirt are present in the fleece. The clean fiber possesses a natural "crimp" or waviness. The external layer is made up of a large number of fine scales attached to the cortex for about two-thirds their length and all pointing towards the fiber tip [Fig. 7 (a) and (b)]. This results in a distinct difference in frictional resistance in traveling to or from the root-end, amounting in some wools to as much as 60%. Wool differs from cotton in being much more extensible and hygroscopic, taking up 30% of water without appearing wet; its strength, however, falls continuously with increasing moisture content.* Chemically, wool consists of a protein (p. 380), keratin, exceptional in containing about 3.5% sulfur combined in cystine residues (p. 160) having disulfide linkages. It has an X-ray pattern of identity period 5.1 A along the fiber axis, in common with many other epidermal growths. Like other proteins, wool is amphoteric, having an isoelectric region at about pH 4.9.

* The differences of the textile fibers in behavior toward moisture are important. Thus, linen absorbs liquid water more rapidly than cotton, perhaps due to its higher pectin content. The difference between cotton and wool is still more striking, the latter even when water-soaked preserving its fiber resilience to a far higher degree. The insulating capacity of a fabric depends predominantly on the air effectively entrapped in the voids between the fibers. Comparing two similar fabrics, the water-soaked woolen one tends to hold its open structure, retain its heat-insulating capacity, and retard the evaporation of water with a consequent cooling effect, while the cotton fabric is prone to collapse when wet and cling to the underlying surface. By a suitable manufacturing technique fabrics such as blankets can be produced from the cheaper cotton fiber with insulating capacity equal to or greater than that of wool. Though such fabrics on washing tend to deteriorate in this respect more rapidly than does wool, they can hold their advantage for considerable periods when skillfully fabricated. The superiority of wool can be secured in large degree by the incorporation of a limited percentage of woolen fibers in the total, giving importance to the use of cotton-woolen mixtures or union goods.

Dry wool can be stretched about 20% of its original length, but, by steaming or soaking in water at 90° C., 100% extension can be obtained. The same maximum extension can be secured at room temperature in, *e.g.*, 1% aqueous sodium hydroxide. Furthermore, the magnitude of the ultimate extension is independent of the method of stretching or load employed. If the stretched wool be examined with the X-ray, its pattern (Fig. 8) changes from that of α-keratin to that of a new form, β-keratin, in which the identity spacing is 6.8 Å, very close to the value, 7.0 Å, found in silk fibroin (cf. Fig. 4, p. 162). Provided the load is almost immediately removed, the elongation is reversible, α-keratin being again obtained. If the fiber be dried before the load is removed, contraction does not occur in a perfectly dry atmosphere; the presence of water is essential for either contraction or expansion. After intermediate but short times of steam-
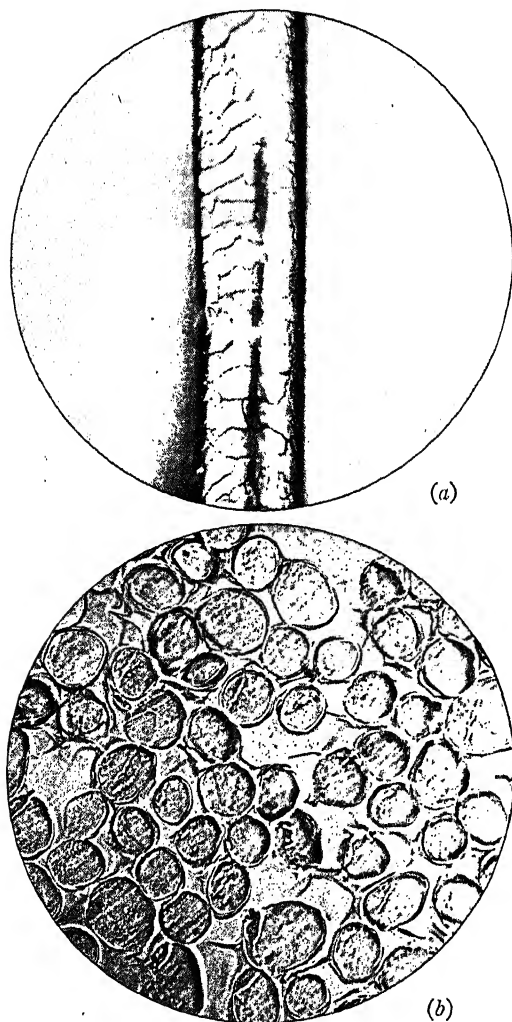
Fig. 7. Photomicrographs of Wool (× 300).*

(*a*) Longitudinal section
(*b*) Cross section

* Courtesy of Prof. E. R. Schwarz, Textile Laboratory, Massachusetts Institute of Technology.

ing under load, release of load causes supercontraction to 70% or less of the initial fiber length. With longer steaming times this phenomenon becomes less marked, the fiber ultimately showing permanent set. The maximum supercontraction is attained more rapidly the higher the stretching temperature but
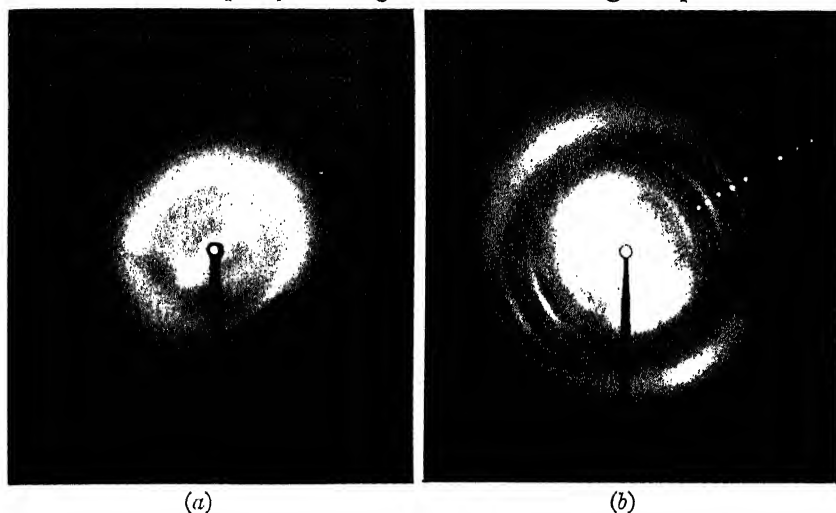
(a)   (b)
α-keratin, unstretched wool, hair, horn   β-keratin, stretched wool, hair, feathers

Fig. 8.   X-ray Diagrams of Stretched and Unstretched Wool.*

its magnitude is independent of the conditions used. Furthermore, the work required to stretch the wool decreases steadily with rise in temperature.

Additional clues as to the changes occurring are also afforded by chemical behavior. Evolution of hydrogen sulfide from wool saturated with water can be detected at moderate temperatures. Reagents, e.g., sodium sulfide and sodium bisulfite, which are known to react with the disulfide link,

$$R-S-S-R \xrightarrow{\text{Na}_2\text{S}} 2R-SNa$$

$$R-S-S-R \xrightarrow{\text{NaHSO}_3} R-SNa + R-S-SO_3H$$
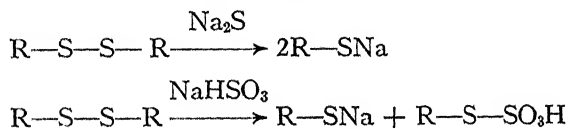
cause the wool fiber to contract. Boiling neutral sodium sulfite solution, on the other hand, causes no contraction of the fiber

* Courtesy of Dr. G. L. Clark, Department of Chemistry, University of Illinois.

although it is known to react with disulfides according to the
scheme:

$$R—S—S—R \xrightarrow{Na_2SO_3} R—SNa + R—S—SO_3Na,$$

but if the wool be previously deaminated by nitrous acid, there
is over 30% contraction in ten minutes.

These facts have been explained by Astbury and Speakman,
assuming for wool a chain structure, cross-linked by cystine and
salt bridges as shown diagrammatically below: *

* The electronic bond between base and acid (—COO—NH₃—) is called a "salt" bridge
or linkage.

In α-keratin, the chains are assumed to be folded due to attractions between the carbonyl and imino groupings according to the plan:

$$
\begin{array}{ll}
\text{α—keratin} & \text{β—keratin}
\end{array}
$$

CHR—CO $\longleftarrow$ CO

CO   NH   NH

NH   CO   CHR

5.1Å   CHR   NH   CO

CHR   NH

CHR   RCH

CHR   CO   CO

CO   NH   NH

NH   CO   CHR   3×3.4 =10.2Å.

CHR   NH   CO

NH

RCH
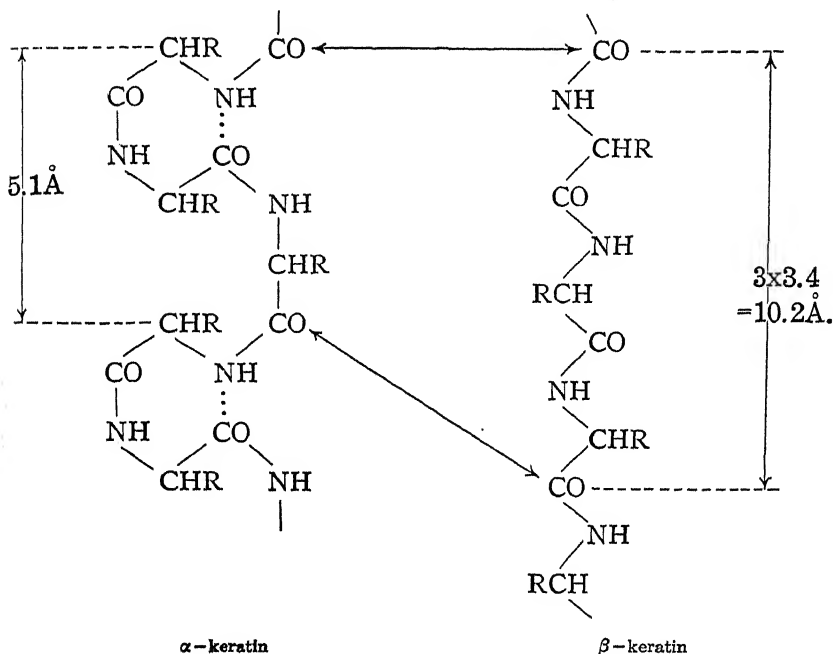
On stretching, the fold can be opened up, yielding β-keratin, although such a change is possible only if the salt linkages are loosened in water or a similar polar liquid. It will be seen that this accounts quantitatively not only for the new identity spacing but for the 100% measured extension.

Supercontraction and permanent set remain to be accounted for. The first change during steaming of a stretched fiber is thought to be gradual breaking of the salt and cystine linkages, the former being sensitive to both low and high pH, the latter to alkaline solutions only:

$$
R\text{—}S\text{—}S\text{—}R \xrightarrow{H_2O} R\text{—}SH + R\text{—}S\text{—}OH.
$$
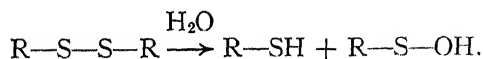
At this stage the chains are completely free and on release of tension, because of elimination of the salt linkages, supercontrac-

tion can occur to a structure even more folded than α-keratin:

$$
\begin{array}{ccccc}
& \text{CHR} & & \text{CO} & & \text{NH} \\
\diagup & & \diagdown & & \diagup & \\
\text{CO} & & \text{NH} & & \text{CHR} & \\
| & & \vdots & & & \\
\text{NH} & & \text{CO} & & \text{CHR} & \\
& \diagdown & & \diagup & \diagdown & \\
& \text{CHR} & & \text{NH} & & \text{CO} \\
& & & \vdots & & | \\
& \text{CHR} & & \text{CO} & & \text{NH} \\
\diagup & & \diagdown & & \diagup & \\
\text{CO} & & \text{NH} & & \text{CHR} & \\
| & & \vdots & & & \\
\text{NH} & & \text{CO} & & \text{CHR} & \\
& \diagdown & & \diagup & \diagdown & \\
& \text{CHR} & & \text{NH} & & \text{CO} \\
& & & & & | \\
\end{array}
$$
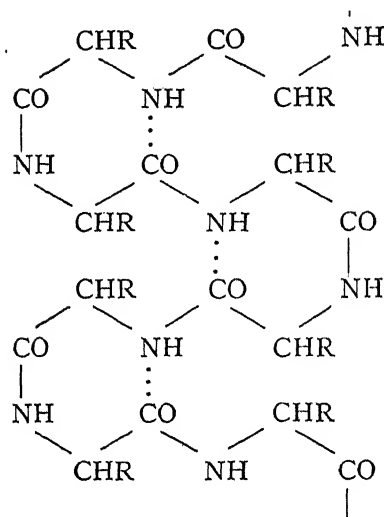
This would have a length approximately 2/3 that of α-keratin in the direction of the fiber axis, agreeing roughly with the shrinkage experimentally found. If the tension is not released, building up of new bridges can now occur, which, increasing in number and strength, finally give a permanent set to the fiber. Both amino groups and sulfur appear to be essential to this rebuilding process, since deaminated wool shows no permanent set even after boiling with water for several hours, and wool treated with baryta water to remove sulfur does likewise. This suggests that permanent set is due to the reaction

$$\text{R---SOH} + \text{R---NH}_2 \rightarrow \text{R---S---NH---R} + \text{H}_2\text{O}$$
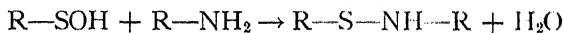
although salt linkages and sulfur bridges probably also occur.*

From the practical point of view one must remember that the fiber, in effect plasticized by moisture, heat, and chemical action, is capable of elongation to double or contraction to 70% of its initial length, after which it can set or "freeze" permanently into either condition. Furthermore, the final set is due

---

* This appears to be the mechanism of permanent waving. The disulfide links are first broken in a steaming process; then, with the hair in the desired new position, fresh cross-links are allowed to form.

to development of chemical cross-linkages of considerable stability. There is even the possibility of linkages between adjacent fibers. These facts are the basis of the felting of wool (see below).

Wool Scouring and Carbonizing.

Wool is freed from grease and dirt by countercurrent washing or scouring. Clearly, pH and temperature must be controlled at each point in the operation to avoid modification of the fibers, and mechanical handling likewise minimized to prevent felting. An alternative method is solvent extraction of the grease with naphtha, after which dirt is readily removed by water washing, the natural soaps of the suint in some cases being sufficient for detergency. The clean fiber is combed before spinning to separate it into the longer fibers which go into worsteds and the shorter ones used for woolens. Wool is frequently not bleached. When this is necessary it is done by padding either in a dilute solution of sulfur dioxide or in a warm, slightly alkaline solution of hydrogen peroxide, sodium perborate, or similar per-salt. "Stoving" in gaseous sulfur dioxide may also be employed. All these reagents attack the cystine linkages of wool and care must be taken to avoid permanent degradation of the fiber. Because complete mechanical removal of cellulosic material, *e.g.*, burrs, is difficult, the last traces are usually eliminated by carbonization, *i.e.*, impregnation with a dilute solution of mineral acid or acid-reacting salt, such as aluminum chloride, followed by drying at elevated temperature, 200 to 250° F. This converts the cellulose to hydrocellulose which readily dusts out, with no injury to the wool. The operation is frequently applied to the woven cloth.

Felting.

The operation of felting, fulling, or milling is probably the oldest technique employed for the production of artificial fabrics. It depends upon the fact that if wool is mechanically manipulated while moist and hot the fibers will interlock and coalesce to a degree that can convert even a loose mass of them into a dense fabric. The operation is applied industrially both to

woven material and to mats of raw fibers (fiber felt). It is particularly important in the finishing of woolens, making it possible to impart to them a density and strength which the shortness of the fiber would otherwise prevent.* The operation may be conducted under heavy hammers (fulling stocks); or, particularly when less drastic felting is desired, e.g., for flannels, in a milling machine, such as is shown in Fig. 9. The cloth, folded double and face-side inwards for protection, travels as an endless belt round and round through a pair of squeeze rolls to a spout, made of oak boards. Here most of the milling occurs, the cloth being gathered in loose folds between the boards. The
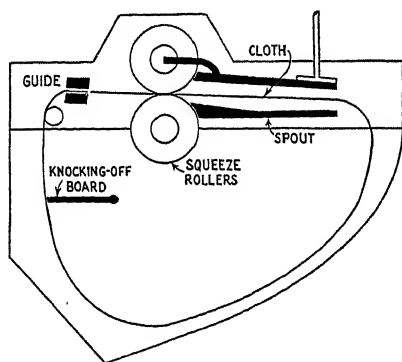
upper board, which can be weighted to secure different degrees of fulling, is moved up and down mechanically (mechanism not shown). The best handle or feel of the finished goods is secured by milling in soapy water at about 45° C., although milling in sulfuric or other acid is sometimes employed as faster. Worsteds

Fig. 9.   Milling Machine.†   may be milled for only a few hours to secure a superficial felting, but many hours, even days, are required for goods such as broad-cloths. After milling, the cloth or felt requires careful washing and is then dried or tentered (p. 490).

The ability to shrink and felt successfully depends upon a number of variables. Mechanical working is necessary but must not be too vigorous; there is an optimum temperature of about 45° C.; the water used should be of low or high pH, preferably about 10. The last is readily obtained by using soap, together with a little alkali carbonate. Different wools vary greatly in felting power, which more or less parallels scaliness.‡

* Unfortunately, felting occurs to at least a limited degree in ordinary laundering. Dry cleaning, i.e., extraction with non-polar organic solvents, has the advantage of avoiding the corresponding shrinkage.

† Scott, E. K., Proc. Text. Inst. 24, 247 (1933).

‡ The interrelation between fiber scales and felting is confirmed by the fact that descaled fibers will not felt.

Picture a mat of wool fibers, "plasticized" by moisture and heat, and submitted to the mechanical action of milling. Under the distortion many of the fibers will elongate and in consequence slip past their neighbors. Because the load is quickly released, the stretched fibers contract. However, these fibers find it difficult to slip back into their original positions over their neighbors, because of the increased friction due to their scales. In other words, the scales tend to allow one way slip only. Consequently, the contraction of the fibers will pull their neighbors closer together, resulting in shrinkage of the whole fabric and compacting of its structure. Because elongation occurs repeatedly and conditions are chosen for maximum plasticity, a considerable amount of supercontraction of the ultimate fibers is likely to occur, making for still greater shrinkage and density. As a result, one can produce by felting even from short fibered wools fabrics of remarkable strength.

The strains of spinning and weaving tend to develop a more or less permanent set in the dry fabric but lead to shrinkage on washing, due to the elastic recovery of the wet fibers. To eliminate this tendency and improve appearance, the goods are often submitted to crabbing or blowing, treatments with steam or hot water or both under slight tension. Presumably, mechanical readjustment of the fibers is secured by a formation of new cross-linkages corresponding to conditions of minimum stretch.

Unshrinkable Wools.

Unshrinkable wools are prepared by chlorination of the finished goods. The stock, impregnated with about 2% aqueous hydrochloric acid, is worked in calcium hypochlorite solution (1.5 to 2° Tw.) for about twenty minutes, rinsed, and treated with an antichlor such as bisulfite or sulfur dioxide. The goods are finally run through very dilute soda ash solution and washed. Fibers thus chlorinated show under the microscope that the scales have been disintegrated. The capacity for further felting is in fact largely eliminated, confirming the significance of the part played by the scales in the milling. The chlorinated goods are often harsh and difficult to dye. The mechanism of chlorina-

tion, particularly that of the attack on the scales themselves, is unknown, although chloramines undoubtedly are formed.

Artificial Wool.

A recent development is a wool substitute prepared from milk casein dissolved in alkali, spun into acid like rayon (p. 363), and the filaments hardened by subsequent treatment in formaldehyde. Artificial wool (lanital) has no scale structure and possesses less extensibility and strength than wool (Table I).

<div align="center">SILK</div>

Silk is the fiber obtained by unwinding the cocoon of the silk worm after softening in hot water. Microscopic inspection shows that the fiber or bave consists of two ultimate threads (brins) lying side by side, cemented together by a gummy material which surrounds each brin. The threads consist of the protein, fibroin, and the gum of an entirely different protein, sericin, the proportion varying but averaging 25% sericin and 75% fibroin.

Sericin dissolves slowly in hot water but disperses readily in hot soap or sodium carbonate solution. Fibroin is insoluble in water and attacked very slowly by dilute alkali. It is, however, dissolved readily by concentrated mineral acids. Hydrolysis of fibroin yields over 60% of glycine and alanine (Table III, p. 161). If one assumes the following relatively simple structure:

$$\begin{array}{ccccccc}
 & & CH_3 & & & & CH_3 \\
 & & | & & & & | \\
NH & CO & CH & NH & CO & CH & NH \\
CO & CH_2 & NH & CO & CH_2 & NH & CO
\end{array}$$
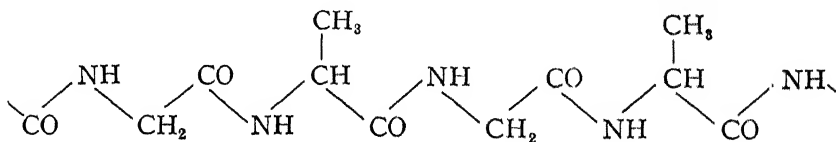
one would anticipate an average chain length per amino acid residue of 3.5 Å. X-ray examination (see p. 162) shows the fiber pattern with a spacing of exactly this magnitude. This structure corresponds to the physical properties of the fiber, which shows but little increase in extensibility on wetting, but a high tensile strength—considerably greater than that of wool. The facts indicate that the silk fiber is made up of relatively straight polypeptide chains arranged side by side, extension occurring
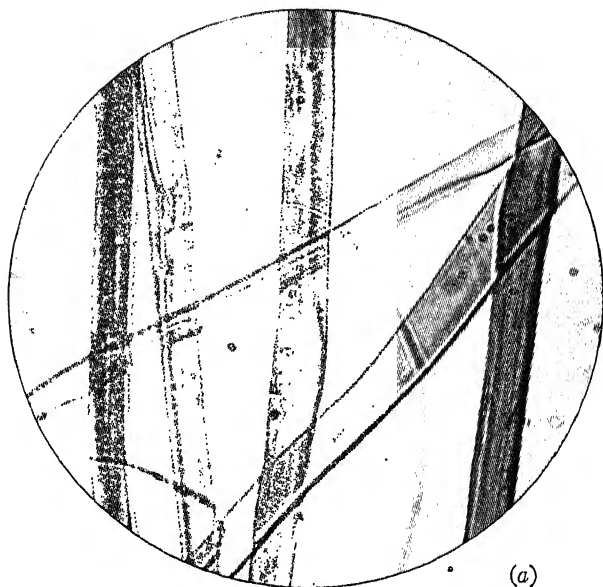
FIG. 10. Photomicrographs of Degummed Silk Fiber.*

(a) Longitudinal section (× 390)
(b) Cross section (× 650)

* Courtesy of Prof. E. R. Schwarz, Textile Laboratory, Massachusetts Institute of Technology.

TABLE II. CHEMICAL REACTIONS

Fiber	Water	Heat	Concentrated Sulfuric Acid	Concentrated Nitric Acid	Concentrated Hydrochloric Acid
Cotton (viscose cuprammonium linen)	Stable. Molds and bacteria cause rotting	Oxycellulose formed in air. Scorches at 150° C.	Dissolves on heating	Mercerizing action followed by nitration	Dissolves very slowly
Wool	Gradually reduces tensile strength. Permanent set in boiling water	Turns brown at 140° C., scorches at 200° C.	Slowly dissolves on heating	Stains yellow (xanthoproteic reaction), dissolves slowly	Nearly unaffected
Silk	Boiling water reduces both luster and strength	Scorches and loses strength about 140° C.	Dissolves with yellow color within 2 minutes	Stains yellow, dissolves quickly	Dissolves in 30% and above

only by slip. The relative absence of side chains brings the polypeptide chains close together, increasing cohesion and fiber strength.

Degumming and Weighting.

The sericin is removed by taking advantage of its solubility in hot soap solution, leaving the fibers as shown in Figs. 10 (a) and (b). The "boiling-off" is usually performed by working the skeins up and down in an 8 to 10% soap solution just below the boiling point. Some sericin may be left undissolved, producing "souple" or "ecru" silks. Soft water, the use of a mild soap (e.g., olive oil), and careful pH control at about 10 to 10.5 are desirable. Even so, losses in strength up to 40 to 50% can result.*

Bleaching is usually effected with sulfur dioxide or bisulfites although, as with wool, peroxides may be used. Color permanence requires good degumming. Impregnating with a half per

* Silk can be degummed by fermentation, a process sometimes employed for waste silk or floss.

OF THE PRINCIPAL TEXTILE FIBERS

Dilute Mineral Acids	Organic Acids	Concentrated Sodium Hydroxide	Boiling 2% Sodium Hydroxide	Chlorine	Sulfur Dioxide
Tenders (hydro-cellulose)	No action	Swells, mercerizes	No action in absence of air	Unaffected at low concentrations under favorable conditions	Tenders if oxygen present
Very little action	No action, readily absorbs	Dissolves	Dissolves	Destroys scales, in larger amounts fiber is badly damaged	Unaffected
Boiling causes gradual loss of luster and strength	Increases luster and "scroop"	Dissolves	Dissolves more slowly than wool	Destroys	Unaffected

cent of acetic or tartaric acid, squeezing and drying without washing, imparts rustle (scroop).

Silk is extensively weighted (sometimes as much as 400%) because this confers improved luster and feel. The process is accompanied by an increase in cross section of the fiber roughly proportional to the gain in weight, but, if carried too far, it ultimately results in brittleness. While materials such as sugar, and thorium and titanium salts are used, the most important weighting materials are probably tin salts. After impregnation in stannic chloride solution and dipping into water, the material deposits in the fiber by hydrolysis as a basic salt and is fixed by padding in some form of alkali, e.g., sodium phosphate, sodium silicate, etc. Careful elimination of chlorides by washing is important, since otherwise in the presence of perspiration the fiber is tendered.

DYEING

Colored materials which can be taken up directly by textile fibers are known as dyes. Formerly obtained exclusively from

natural sources, they are now predominantly synthetic in origin. They differ widely in type. A first group, the direct or substantive dyes, consists of water-soluble colors which are taken up selectively by various fibers, including both cellulose and wool. Their action is substantially reversible and apparently adsorptive in character. The next two classifications are the acid and basic dyes, again water-soluble, both classes being picked up preferentially by both wool and silk, evidently by chemical reaction with the amphoteric fiber proteins. The third group consists of completely insoluble colors, which can be deposited, apparently within the fiber itself, either by synthesis in situ or by chemical conversion into soluble forms from which the color body can be regenerated after introduction into the fiber. The first sub-group are the ingrain colors (ice colors), the second, the vat colors. It seems not unlikely that the so-called sulfur colors and certain mineral colors such as Prussian blue (which can be precipitated in the fiber) also belong in this classification. In any case, because the color body itself is insoluble and deposited within the fiber, even though just below its surface,* this whole group of colors gives dyeings of extraordinary fastness, particularly to mechanical agencies such as washing and the like. Finally, there is a group of colors, possessing little direct affinity for the fiber, which none the less can be anchored to it by the use of a third substance with affinity for both fiber and dyestuff. These are the mordant colors. Because each type requires a particular technique of application, they will be discussed separately.

Organic dyestuffs are molecularly complex, but always contain groups which confer color, the so-called chromophores—the azo group ($-N=N-$), the nitro group ($-NO_2$), the nitroso group ($-NO$), and the o- and p-quinonoid groupings $\left(\text{⬡} \text{ and } =\text{⬡}= \right)$. Such a group converts a molecule into a chromogen capable of imparting color, but does not give it dyeing properties. For the latter, auxochromes must be introduced, e.g., groups such as $-OH$, $-NH_2$, $-N(CH_3)_2$, $-HSO_3$, etc. Thus, azobenzene is not a dye but many dyes are derived

---

* Rate of diffusion of the solutes into the fiber is often very slow.

from it by the introduction of auxochromes, *e.g.*,

$$\langle \rangle\!-\!N\!=\!N\!-\!\langle \rangle\!-\!NH_2$$
$$NH_2$$

chrysoidine
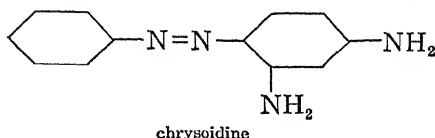
In general, the more complex a molecule and the greater the number of its chromophores, the darker its color.

Owing to the differences in dyeing technique for various colors and to the forms in which textiles may be dyed, *e.g.*, loose, yarn, warp, piece goods, etc., the machinery employed is extremely varied. Thorough and even impregnation of the goods must be provided by continuous relative motion of the dyebath and the material.

The cheapest method of dyeing is submerging the loose fibers in the dye liquor and circulating it through them in an open vat or compressed in a closed vessel with a false bottom, but the stock must be segregated through all subsequent operations. Hanks are dyed suspended from wooden rods stretched across an open, rectangular wooden box or dyebeck, often lined with monel metal or copper. A perforated steam pipe is placed at the bottom, the hanks being rotated on the rods or the liquor circulated through them mechanically. In so-called machine dyeing, on perforated spools (metal, wood, or cardboard) the yarn is wound with relatively large pitch, thus giving an open structure. The dye liquor is recirculated radially through the spools; direction of flow is reversed when necessary to get even dyeing. Penetration of the dye into the yarn and even into the fiber itself is often promoted by certain emulsoidal colloids in the dye bath (e.g., gum tragacanth, although various synthetic products are generally used). The method reduces labor costs. Positive circulation of large quantities of dye liquor through the stock at high velocities, obviously tends to help redistribution of any dyestuff unevenly deposited in the initial stages of treatment.

Piece goods may be dyed in open-width on a jigger, a V-shaped wooden or metal-lined trough with bottom steam. Loose guide rollers are placed near the bottom and top, the cloth being wound back and forth through the dye bath from one large roller to another placed above the trough. Generally about four to six passes are given. A light or medium shade cloth may be dyed by a single pass through a padding machine, dyestuff solution being added continuously. A winch machine is used for knitted and similar goods, which cannot be placed under tension. It consists of a large rotating roll above the dye bath; this serves to keep the goods, sewn into an endless band about it, moving slowly in the liquor. Hosiery and similar materials may be held in an inner, rotating, perforated drum submerged in the bath.

Because the direct or substantive dyestuffs dye both cotton and wool, or silk, without mordant, they are largely used for union goods (mixed wool and cotton fabrics). Chemically they are usually water-soluble sodium salts of sulfonic acids containing one or more azo linkages.* Typical examples are:

chlorazol sky blue FF

benzopurpurin 4B
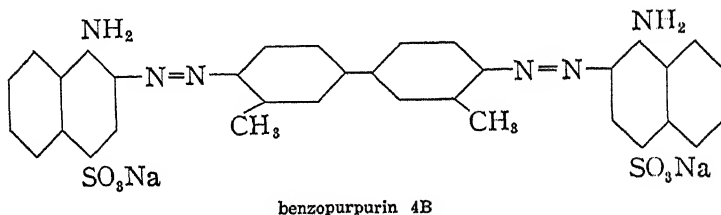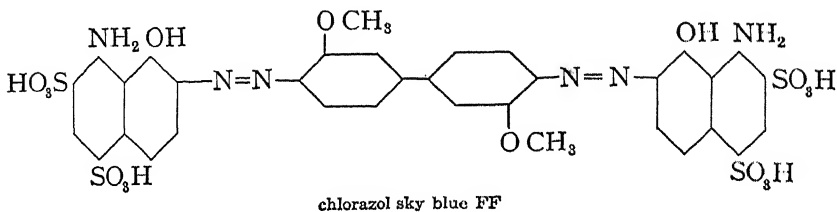
The material is treated with a solution of the dyestuff at or near the boil. Light shades are easy to obtain; but in order to exhaust the bath, "assistants" are added; this is particularly desirable when it is necessary to secure darker or heavier dyeings. These consist of sodium sulfate or chloride (10 to 20% on the weight of cloth). For wool and silk, 1 to 5% acetic acid is also usually added. Direct dyes give full, fairly bright shades, but lack fastness to washing. If the dyestuff contains free hydroxyl groups, fastness may be improved by after-treatment with metallic salts, e.g., copper sulfate, sodium bichromate, etc. Formaldehyde sometimes improves fastness.†

Acid dyes are mainly soluble sodium salts of carboxylic or sulfonic acid compounds, frequently containing an azo group as chromophore. Some are sulfonated basic dyes. Typical ex-

---

* These linkages may be broken by reducing agents, thus destroying the dye. This affords a convenient method of "stripping" or removing unsatisfactory shades from the fabric.

† If the dye contains an amino group, it may be diazotized on the fiber and developed in a solution of a phenol (cf. Ice colors, p. 510). Diazotized p-nitraniline may also be "coupled" with the dyed fiber.

amples are:
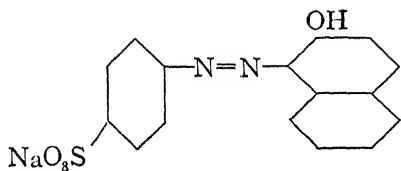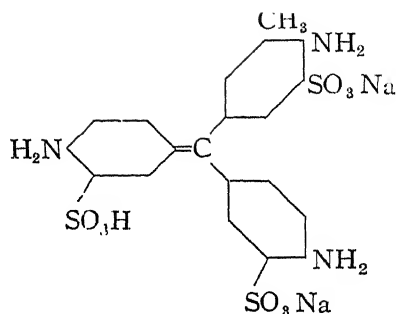
orange II   acid magenta

In the presence of acid they dye wool and silk directly, being used principally for wool and wool-silk unions. For cotton, an alum mordant must be used but even so, the shades obtained are very loose to washing. On wool and silk fastness is moderately good to washing and excellent to light. In dyeing, about 10 to 25% sodium sulfate and 1 to 5% sulfuric acid are added as assistants, the goods being entered at 140° F. and the temperature raised slowly to the boiling point. Wool is frequently dyed before milling, particularly where heavy felts which would impede the penetration of the dye liquor are to be produced; here dyestuffs fast to milling must be chosen. For silk, boiled-off liquor is usually added to the dye-bath to maintain luster and assist level dyeing. In some cases the dyeing can be done in the boiling-off operation. Alkalies strip most acid colors completely from dyed wool or silk.

Basic colors, in contrast to the acid dyes, are salts of organic bases, generally containing a quinonoid grouping, *e.g.*,
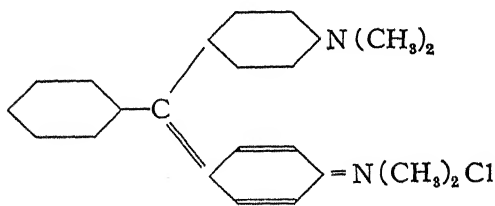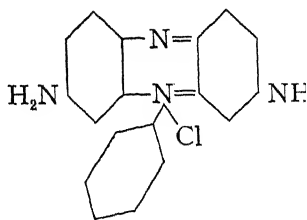
malachite green,   phenosafranine

As a class, they possess great tinctorial power and brilliancy but

little fastness to washing or light; hence, they are little used except on silk. For cotton, a mordant (usually tannic acid, with which basic dyes form insoluble complexes) is required. As might be expected they can often be stripped by boiling hydrochloric acid.

The ice, azoic, or ingrain colors are essentially azo pigments synthesized in situ on the fiber. They are very fast to washing and many can even be boiled in dilute sodium carbonate without injury. In addition, vivid shades can be obtained. They are, however, unsuited for wool and silk. In illustration, cotton is padded in an alkaline solution of β-naphthol together with Turkey Red oil * to assist penetration, and dried in absence of light. The goods are then "developed" in a solution of a diazotized base, e.g., p-nitraniline (yielding para red), generally containing sodium acetate to remove free hydrochloric acid, which would prevent coupling. The color forms immediately.

$$R.NH_2 + HNO_2 + HCl \longrightarrow R-N=NCl + 2H_2O$$
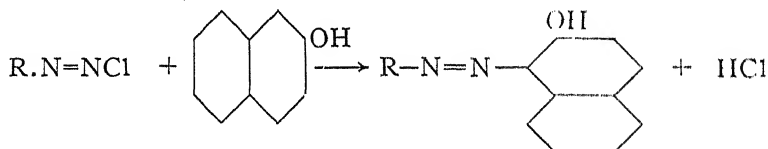
Dyeing is usually followed by treating with hot dilute soap solution, "soaping," to remove loosely adhering color † and improve brightness. Because β-naphthol has no affinity for the fiber, evenness of shade is difficult to attain; this has been overcome to some extent by the use of naphthol A.S. (anilide of β-oxynaphthoic acid) and its analogues, which have a definite affinity for cellulose, making it unnecessary to dry the naphtholated material before developing. Stabilized diazonium compounds (double salts with inorganic chlorides or sulfates) may also be used.

The vat dyestuffs are among the oldest of the natural dyes.‡ They are all characterized by one property—while themselves insoluble, treatment with a reducing agent gives leuco com-

---

* Sulfonated castor oil converted into the sodium, potassium, or ammonium salt.

† It is difficult to prevent at least traces of the naphthol from diffusing out of the fiber into the developing bath and precipitating there.

‡ Indigo and Tyrian Purple.

pounds soluble in alkalies. Thus, for indigo,

indigo   reducing agent / NaOH   soluble leuco base
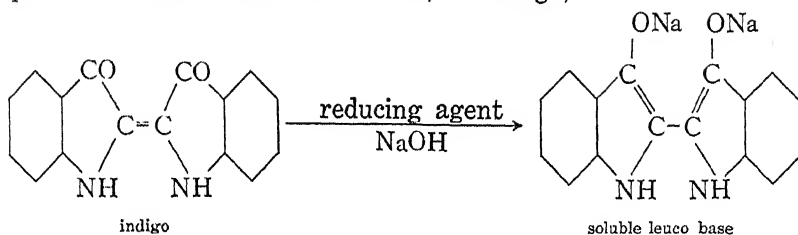
Anthraquinone also gives rise to many dyestuffs of this type, *e.g.*,

algol pink R
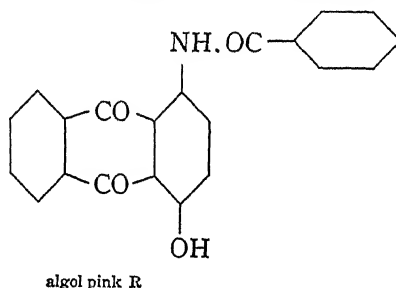
Vat dyestuffs are brought into solution by reduction with sodium hydroxide and hydrosulfite. The goods are dyed in this at room temperature or above, as far as possible keeping them below the surface of the dye bath. After squeezing, exposure to the air causes the leuco compound to be oxidized back and precipitated on or within the fiber as the original, insoluble pigment. As in dyeing with azoic colors, hot soaping follows. Vat colors have good fastness to washing, light, alkalies, and acids, being used principally for shirtings, dress materials, household fabrics, etc. Application to wool and silk is difficult because of the highly alkaline bath required; in addition, dyestuffs are available which are almost equally fast and more easily applied.

Little is known of the chemical constitution of the sulfur dyestuffs, many of which are formed by heating organic bases with sulfur. They are insoluble in water but are brought into solution by warming with a concentrated sodium sulfide-sodium carbonate solution, subsequently adding this to the dyebath, from which dyeing is carried on much as with the direct colors. Because of the high alkalinity and sulfide requirements, they are little used for wool and silk but, in spite of low tinctorial power,

are largely used on cotton, since they are cheap and have good fastness to light and washing. Owing to sulfuric acid formation, goods dyed with sulfur colors are particularly likely to tender on storage. This can be minimized by after-treatment with sodium bichromate and acetic acid, which presumably oxidizes immediately any loosely combined sulfur, allowing it to be washed out.

Mineral colors have practically lost their importance but are occasionally employed because they are fast to light and washing. In general, an insoluble colored salt, *e.g.*, lead chromate, Prussian blue, iron oxide, is precipitated in the fiber, where it is mechanically held.

The mordant colors, as the name indicates, are dyed with the aid of an intermediate material, usually a chromium or other trivalent metallic salt. All the dyes appear to form insoluble compounds or lakes with these salts, owing to the presence of hydroxyl or carboxyl groups. Furthermore, two groups, one of which must be hydroxyl, and the other hydroxyl, carboxyl, nitroso, or azo, appear to be essential. Typical dyes of this class are:
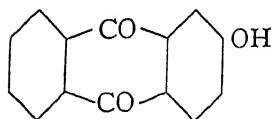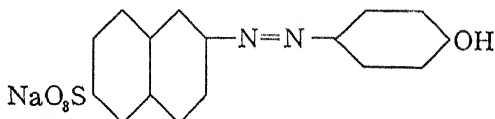
alizarine   anthracene chrome yellow

Logwood is the only natural dyestuff of this class now used. The mordant colors are generally fast and good on wool. Chrome mordants are applied with sodium bichromate, either with the addition of a reducing agent (cream of tartar, lactic or oxalic acid) turning the wool fiber green, or with sulfuric acid producing an unreduced mordant, which should be dyed immediately. The mordanted goods are entered hot and dyed at the boil in a solution of the dyestuff, often containing calcium acetate as an assistant.

Cellulose Acetate Dyeing.

Cellulose acetate lacks affinity for all common dyes except basic ones and a few isolated, individual dyes. The problem has

been solved by the introduction of two new classes of dyes. Insoluble amino-azo compounds, for which the acetate has an affinity, are rendered water-soluble by formaldehyde-bisulfite. The methyl ω-sulfonic acids produced gradually hydrolyze during dyeing and the regenerated, insoluble azo bodies are taken up by the fiber.   A typical member of this class is Ionamine A.S.

$$(H_3C)_2N\langle\rangle-N{=}N-\langle\rangle NH.CH_2.SO_3Na.$$
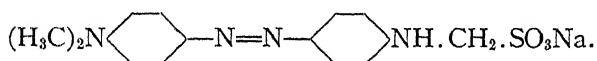
The color can be diazotized and developed to the ingrain type, as described for certain direct cotton dyes.   The second class of dyestuffs is prepared by colloidally dispersing aminoanthraquinone derivatives by grinding with Turkey Red oil or soap. Such materials appear to be taken up in solid solution by the acetate fiber.

Theories of Dyeing.

With substantive dyes on cotton, a chemical mechanism of adsorption seems unlikely.   Confusion has been caused in the past by failure to distinguish between rate of dyeing * and final equilibrium, but when equilibrium is attained it is perfectly reversible, *i.e.*, approachable from either side.   Furthermore, the amount of dye absorbed at equilibrium is given approximately by either the Freundlich isotherm or the Langmuir equation (p. 85).   For many substantive dyes, the amount is surprisingly independent of hydrogen-ion concentration (cf. wool and acid and basic dyes).   In common with most adsorptive processes, temperature has a considerable effect upon the equilibrium (Fig. 11).   That adsorption increases in swollen cellulose, *e.g.*, rayon or mercerized cotton, is shown by Fig. 12.

The kinetics of the dyeing process have been investigated by Neale and his collaborators, using regenerated viscose or Cellophane as a form of cellulose, in which it is actually possible to measure the rate of diffusion of the dye.   On immersing a Cellophane sheet into the dyebath, the sides at once become

---

* This may be astonishingly different for different dyestuffs or even for the same dyestuff under varying conditions.   Benzopurpurin 4B (p. 508) takes four seconds using 25% salt at 90° C. to reach equilibrium and twenty-one hours at 30° C. with 0.2% salt, the first rate being 18,900 times faster than the second.

colored, the color gradually penetrating into the interior of the sheet until it becomes uniform. It is found that, to a first approximation, the results can be expressed by Fick's diffusion law,

$$\frac{ds}{dt} = -k\frac{dc}{dx},$$

where ds/dt is the rate of diffusion at any point in the cellulose at which the concentration gradient is dc/dx.

1. Fast Red K on Cotton
(0.05 gm./liter, NaCl 5 gms./liter.)
2. Fast Red K on Cellophane
(0.05 gm./liter, NaCl 5 gms./liter.)
3. Heliotrope 2B on Cotton
(0.05 gm./liter, NaCl 5 gms./liter.)
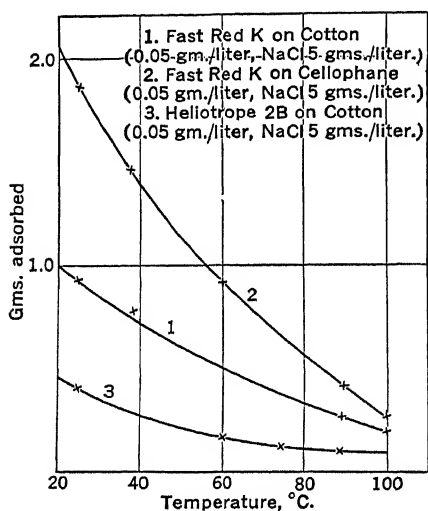
FIG. 11. Effect of Temperature upon the Adsorption of Direct Dyestuffs.*

Nearly all direct dyestuffs are capable of dissociation as colloidal electrolytes. Many of them are ampholytes, e.g., in Chlorazol Sky Blue FF (p. 508) ionization of both amino and sulfonic acid groups is possible. Aggregation in aqueous solution of these large dyestuff molecules into micelles of colloidal dimensions is often found to occur, as shown by the Tyndall beam in the ultramicroscope, streaming double refraction (p. 134), osmotic measurements, and, in particular, ultrafiltration (p. 137) and diffusion experiments. Neutral salts generally increase agglomeration, which in turn tends to increase adsorption on the fiber. The possible magnitude of this effect is illustrated in Fig. 12, where, because of the range covered, the data are plotted logarithmically. It is seen that, whereas in oxycellulose and Cellophane dye adsorption is approximately proportional to salt concentration, in cotton and mercerized fibers the dye held increases less rapidly, roughly as the 0.6 power of the salt. The low affinity of oxycellulose (0.3 gms. oxygen per 100 gms.

* Data from GARVIE, W. M., GRIFFITHS, L. H., and NEALE, S. M., *Trans. Faraday Soc.* **30**, 271 (1934).

cellulose) for the dye agrees with general dyehouse experience. Agglomeration is decreased by rise in temperature.†

The faster a dyestuff the more difficult it is to secure even dyeing. If the yarn or cloth absorbs an excess of dye at a given
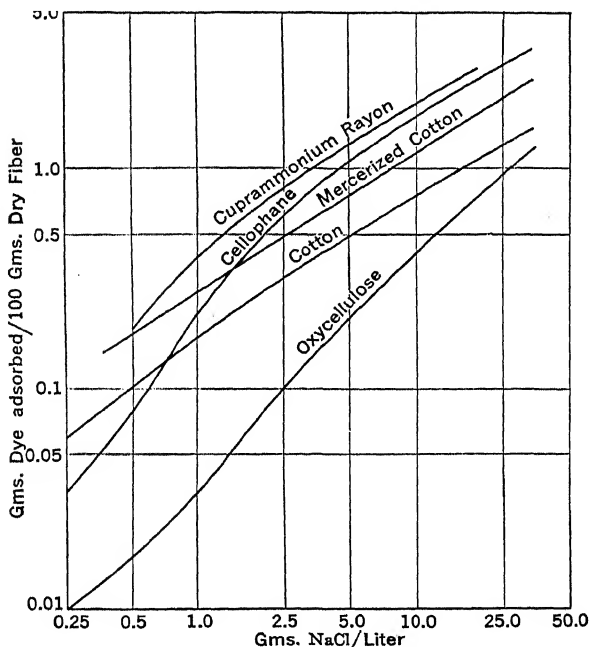
FIG. 12. Adsorption of Chlorazol Sky Blue FF by Cellulose at 90° C. from Dye Solution of 0.05 Grams per Liter.*

* HANSON, J., NEALE, S. M., and STRINGFELLOW, W. A., *Trans. Far. Soc.* 31, 1718 (1935).

† The effect of increasing molecular weight of solute, equivalent to particle size (see p. 102), in increasing adsorption on a solid at low concentrations is illustrated by the data of Table VI on page 91. However, molecular aggregation slows down the Brownian movement (see pp. 99, 109), reducing the rate of diffusion (p. 135) of the dye molecules. Surprisingly, the diffusion coefficient of the dyestuff, at least through Cellophane, rises to a maximum and then falls as salt concentration is increased. The reason for this anomaly is unknown, but the facts indicate that the phenomena are extremely complex.

It has been suggested that in the case of those direct dyestuffs which show little or no adsorption on cotton unless salts are present, the function of the latter is to lower the ζ-potential (p. 189) of the negatively charged cellulose so that the anions of the dyestuff can approach. At low salt concentrations salts of divalent metals are more effective in causing adsorption than are those of monovalent. Aluminum sulfate, however, is less effective than sodium chloride, possibly because it coagulates the dyestuffs. Figure 11 indicates that at high salt concentrations the more swellable fibers give greater color adsorption, due perhaps to greater accessibility of the large dyestuff aggregates to the fiber interior. The relatively low adsorption on regenerated cellulose at low salt concentrations may represent a reduced affinity between color and fiber due to a degradation analogous to that of oxycellulose formation (p. 154).

point this cannot redistribute. Despite its effect on rate of diffusion, degree of dispersion of dyestuff in the bath has relatively little effect on evenness of dyeing, because proper agitation distributes the dyestuff by convection, diffusion playing no essential part. Since penetration of the dyestuff into the fiber occurs by diffusion only, molecular dispersion is important to secure uniformity of distribution throughout the fiber, colloidally dispersed dyestuffs resisting fiber penetration (see Fig. 13). High temperature of the dye bath distends the fiber, opening the spaces between its micelles and promoting diffusion. This

Acid Orange 2G   Solway Blue   Solway Blue
     60.5°   60.5°   97.0°
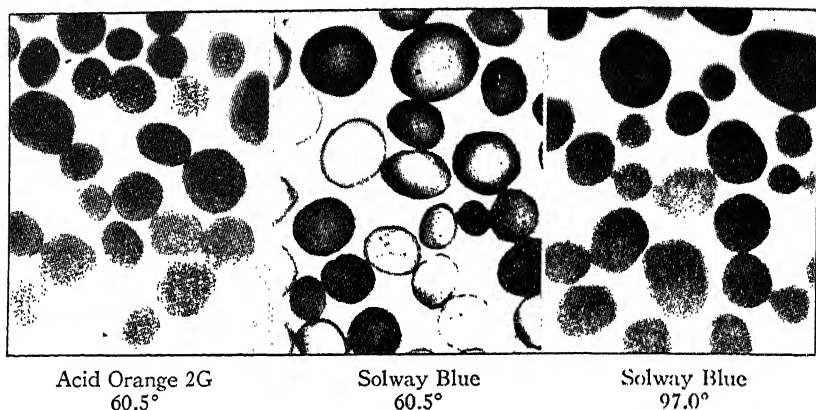
Fig. 13.   Dyeing of Mohair with Acid Orange 2G
and Solway Blue SEN (× 335).*

is outstanding in the case of wool, where increased swelling is possible because of the loosening of the cross-linkages in the molecular structure (see p. 496). The ultimate cause of the affinity of the direct dyestuff for the fiber (substantivity) is uncertain. The dye molecules are presumably held by partial valence forces. For cellulose, long straight dyestuff molecules should most readily fit in between the cellulose chains and micelles and be correspondingly firmly held. Many direct dyes for cellulose have this structure, but other factors, imperfectly understood, undoubtedly play a major part.

In contrast to direct dyes on cotton, the adsorption of acid dyestuffs by wool is found to be markedly dependent upon the

* SPEAKMAN, J. B., and SMITH, S. G., *J. Soc. Dyers and Colourists* **52**, 121 (1936).

hydrogen-ion concentration of the dyebath, and increases rapidly with decrease in pH. Furthermore, the amount of dyestuff taken up is practically independent of the acid added, being a function of pH only. This parallels the behavior of wool and other ampholytes toward ordinary acids (p. 207). Indeed, three distinct stages have been distinguished in wool dyeing with acid colors: rapid combination of the wool with the mineral acid present in the dyebath, which, indicating wool by W—NH$_2$, may be written,

$$W—NH_2 + H_2SO_4 \rightarrow W—NH_2 . H_2SO_4;$$
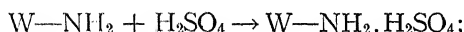
diffusion of dyestuffs into the wool fiber; and metathesis between the wool salt and the dyestuff, or, calling the dye acid HA,

$$W . NH_2 . H_2SO_4 + HA \rightarrow W—NH_2 . HA + H_2SO_4.$$
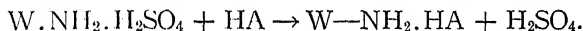
The final concentrations are governed by the Donnan equilibrium (p. 212). Such a mechanism is confirmed by the fact that acid dyes are adsorbed on wool at high dyestuff concentrations in the ratio of their molecular weights. Deaminated wool also shows decreased affinity for acid dyestuffs and the acid combining power of wool decreases on dyeing.

Speakman showed that diffusion apparently controls the rate of adsorption of Acid Orange 2G by human hair. The difference between ordinary and descaled hair is noteworthy, the scales apparently offering considerable resistance to the penetration of dye molecules. This may account for the increased rate of dyeing of chlorinated and damaged wool fibers. Temperature plays a part by affecting the swelling of the wool fiber. Figure 13 shows photomicrographs of mohair dyed with Acid Orange 2G at 60.5° C. and with Solway Blue SEN at 60.5 and 97° C., the former dyestuff being almost molecularly dispersed, while the latter is in colloidal solution. Wool can be dyed satisfactorily with molecularly dispersed dyes either at low or high temperatures, but to obtain proper penetration of the fiber, colloidal dyes should be used only at high temperatures where swelling of the fibers can take place with penetration by the dye micelles. Within a given dyestuff class, colloidal colors on wool are superior in fastness to molecularly dispersed ones, presumably because they are entrapped more firmly in the fiber.

The behavior of basic dyes toward wool is analogous to that of acid colors and requires no separate discussion.

Silk behaves toward dyestuffs much as does wool, undoubtedly due to their common protein structure. With silk, appearance is often more important than fastness, a fact which profoundly modifies dyeing technique. Thus, both acid and basic dyes are employed, the latter particularly for brilliance. Examination of the fiber indicates that the dyestuff is distributed almost perfectly throughout, with a homogeneity which, together with the smoothness of the fiber, gives regularity of reflection and refraction that greatly enhances brilliance.

The chemical technique of application of vat, ingrain, and sulfur colors has already been described. The mechanism by which they are held within the fiber would seem to be largely mechanical, independent of the nature of the dyestuff, other than of its insolubility and permanence. However, the dispersability of the dyestuff and its affinity for the fiber in the different stages of dyeing obviously modify the methods of application and the distribution of the ultimate color on the fiber. These effects in general parallel those discussed under the other types of colors.

BIBLIOGRAPHY

ASTBURY, W. T., *Fundamentals of Fibre Structure*, Oxford University Press, London (1933).

EMMOTT AND CO., *Textile Manufacturers' Yearbook*, Manchester, England (1938–40).

VON GEORGIEVICS, G., *Chemical Technology of Textile Fibers*, Trans. by Salter, C., Scott, Greenwood and Son, London (1920).

HAVEN, G. B., *Industrial Fabrics*, Wellington Sears Co. (1941).

HERZOG, R. O., *Technologie der Textilfasen*, Springer, Berlin (Various volumes).

HESS, K., *Die Chemie der Zellulose*, Akad. Verlag. m.b.h., Leipzig (1928).

HÜNLICH, R., *Textile Fibers and Materials*, Skinner and Co., Ltd., London (1939).

Symposium: New Textile Fibres, Fabrics and Finishes, *Ind. Eng. Chem.*, **32**, 1543–71 (1940).

VALKÒ, E., *J. Soc. Dyers and Colorists*, **55**, 173 (1939).

NOTE: Important material is found in the files of the following periodicals: *Cellulose-Chemie*, Berlin-Dahlen; *J. Soc. Dyers and Colourists*, Yorkshire, England; *J. Textile Inst.*, Manchester, England.

Author Index

Subject Index